



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 174742

TO: Sean McGarry
Location: rem-2d19/2c18
Art Unit: 1635
Wednesday, December 28, 2005

Case Serial Number: 10/774974

From: Deirdre Arnold
Location: Biotech-Chem Library
REM 1A64
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

Please feel free to contact me if you have any questions or would like to amend the search.

Thank you for using STIC services.

Regards,
Deirdre Arnold



This Page Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 00:53:51 ; Search time 13 Seconds

(without alignments)
753.883 Million cell updates/sec

Title: US-10-774-974-2

Perfect score: 7500

Sequence: 1 MMQGYCHRMSPHPCGCRP.....MRWERHQREPDIEDIKK 1374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 54001 seqs, 7132810 residues

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pap.*
- 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pap.*
- 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pap.*
- 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pap.*
- 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pap.*
- 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	276	3.7	948	US-10-523-477-14	Sequence 14, Appl
2	260.5	3.5	245	US-10-793-626-1658	Sequence 1658, Ap
3	221.5	3.0	228	US-11-074-176-136	Sequence 136, App
4	221	2.9	1560	US-11-059-982-1	Sequence 1, Appli
5	211	2.8	244	US-10-467-657-2638	Sequence 2638, Ap
6	207	2.8	795	US-10-770-726-49	Sequence 49, Appl
7	190	2.5	1823	US-10-995-561-988	Sequence 988, App
8	190	2.5	2102	US-10-995-561-990	Sequence 990, App
9	190	2.5	2108	US-10-995-561-989	Sequence 989, App
10	190	2.5	2157	US-10-995-561-991	Sequence 991, App
11	175.5	2.3	175	US-10-821-234-1074	Sequence 1074, Ap
12	174	2.3	1970	US-10-821-234-1641	Sequence 1641, Ap
13	173	2.3	1532	US-10-821-234-914	Sequence 914, App
14	171	2.3	1178	US-11-044-899-29	Sequence 29, Appl
15	170.5	2.3	639	US-10-821-234-907	Sequence 907, App
16	168	2.2	580	US-10-995-561-987	Sequence 987, App
17	168	2.2	5179	US-11-108-172-1068	Sequence 1068, Ap
18	167.5	2.2	1187	US-10-821-234-955	Sequence 955, App
19	163	2.2	2432	US-10-821-234-899	Sequence 899, App
20	162	2.2	572	US-10-821-234-1290	Sequence 1290, Ap
21	155	2.1	186	US-10-508-263-111	Sequence 111, Appl
22	153.5	2.0	438	US-11-140-417-2	Sequence 2, Appli
23	153.5	2.0	4384	US-10-821-234-1120	Sequence 1120, Ap
24	152	2.0	509	US-10-821-234-1093	Sequence 1093, Ap
25	150.5	2.0	456	US-10-477-507A-2	Sequence 2, Appli

26	148.5	2.0	963	6	US-10-467-962B-2	Sequence 2, Appli
27	148	2.0	552	6	US-10-821-234-1022	Sequence 1022, Ap
28	146.5	2.0	793	7	US-11-060-914-2	Sequence 2, Appli
29	146	1.9	564	7	US-11-186-284-199	Sequence 199, App
30	146	1.9	1163	7	US-11-044-899-2	Sequence 2, Appli
31	145	1.9	1163	7	US-11-044-899-30	Sequence 30, Appl
32	144.5	1.9	273	6	US-10-821-234-1203	Sequence 1203, Ap
33	144.5	1.9	432	7	US-11-140-417-4	Sequence 4, Appli
34	143	1.9	676	7	US-11-135-855-28	Sequence 28, Appl
35	143	1.9	717	7	US-11-135-855-29	Sequence 29, Appl
36	142	1.9	356	7	US-11-143-986-8	Sequence 8, Appli
37	142	1.9	356	7	US-11-143-986-9	Sequence 9, Appli
38	141.5	1.9	160	6	US-10-508-263-110	Sequence 110, App
39	141.5	1.9	1516	6	US-10-220-824-8	Sequence 8, Appli
40	141	1.9	2011	7	US-11-080-991-56	Sequence 56, Appl
41	140.5	1.9	1618	6	US-10-984-645-2	Sequence 2, Appli
42	140	1.9	658	6	US-10-821-234-921	Sequence 921, App
43	139.5	1.9	919	6	US-10-821-234-1144	Sequence 1144, Ap
44	139	1.9	530	7	US-11-088-634A-2	Sequence 2, Appli
45	138	1.8	697	6	US-10-821-234-905	Sequence 905, App

ALIGNMENTS

RESULT 1

US-10-523-477-14
; Sequence 14, Application US/10523477
; Publication No. US20050266406A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MAXS AS MODIFIERS OF THE AXIN PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-051C-US
; CURRENT APPLICATION NUMBER: US/10/523,477
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US 60/401,534
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,153
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 948
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-523-477-14

Query Match		3.7%;	Score 276;	DB 6;	Length 948;
Best Local Similarity		23.0%;	Pred. No. 2.7e-10;		
Matches 149;		Conservative 61;	Mismatches 208;	Indels 230;	Gaps 31;
QY	27	ARPSA--PSFRPQNLRLHPPQPPVQYQYEPFSA-----PSTFGNSPAPNLPFR 75			
DB	320	ARPGGRRPGWEHNSN-KLGYLVSPPPQPIRGRSCYRSINRGRHHSERSORTQGPS-LPAT 377			
QY	76	PDVFPPPPMPPSAQGLPPCPPIRPPPHQNMHPPVPPPCFPMPMPMPMPNPPVPGA 135			
DB	378	PVFVPVPPPP-----PLYPPPP-----HTLELP-----PGVPPPPQSPQFP--PGQ 415			
QY	136	PRQQTFFPFWMPPPSPMPPHPPPPVMPQQVN-----YQYPP----- 170			
DB	416	PPAG---YVPPPPGPP-PAPANLSTPWVSSGVQTAHSNTIPTQAPPLSRFFEYRQR 471			
QY	171	-----GYSHNFPFPFSPNS 184			
DB	472	LKEEEKKKSLDEFTNDPAKELMEYKKIKERERSFSRKSYPYSGSSYSRSTYSKRS 531			
QY	185	FQNNPSFLPANNSSSPHRLHPYPLPKASERSRSPERLKYDHRHSDHSHGERH 244			
DB	532	GSTRSRYSRSPSRSHSRYSRSPYP-----RRGRGKSRNY-----RSRSRSHG---YH 578			
QY	245	RLDRERGRSPDRRQDSRYSDYD-RGRTPSRHSYRSRER----- 289			

Db 579 RS-----RSRSPYRRYHRSKRSQAFQSPNKRNVPOQETERYEFNRYREVPPYDMK 633
QY 290 -----ERHRDNRSPSLERSYKKEYRSRSYGLSVVPEPAGCTPELPG 336
Db 634 AYGSRVDFRDPPEKERYE-----WERKYREWYKYYKGYAAGQAPRPSA----- 679
QY 337 IINKTDSWAP-----PLEIVNHRSPSRKKRKRWEKEDKRWSDNOSSG-----KD 381
Db 680 ---NRENFSPEPLPLNIRN---SPFTRGRREDYVGQSHRSRNGNYSPEKLSARDGHNQ 734
QY 382 KNYTSIKKEPEBTMPD-----KNBEEBELKPVWIRCTHSNYSYSSDPM-- 427
Db 735 KONTKSEKESENAPGDKGNKHKKRKRKGESEGFNLPELLETSRK-----SREPTGV 790
QY 428 -DOVGSTVVGTSR-----LRDLYKPEEE-----LGSROKAKAARPPWPPKTKLDELE 478
Db 791 EENKTDSLFLPSRDDATPVRD---EPMDAESITPKSVSEKDKRER---DKPKAKGDKTKR 845
QY 479 SSSESECEDESDSTCSSSDSE-----VFDVIAEIKRKAHPDRL 518
Db 846 KNDGSAVSKENIVKPAKGPOEKVDGVRDLDLNQLKPKRRLRRL 893

RESULT 2

US-10-793-626-1658
; Sequence 1658, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1658
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1658

Query Match 3.5%; Score 260.5; DB 6; Length 245;
Best Local Similarity 30.4%; Pred. No. 4.8e-10;
Matches 70; Conservative 57; Mismatches 88; Indels 15; Gaps 8;
QY 1105 OKLTEPEEAIGVIFTHVRLARAFRTV--GNHLLTGHQNMFLGDSIMOLVATEYL 1162
Db 15 QKPTDRKWSLGLRKNIDLYQAQAFSSHPFNDPMNRLEHNERLEFLGDAVLELTVSYL 74
QY 1163 FIHFDPDHGHLTLRLSSLVNNRTQAKVAELGMQEVATINDTKRPVGLRTK--TLAD 1220
Db 75 FDRHPLPGLNLTGMATKATICEPSLVIFANKINLELILG-KGEKTKGRTRPSLISA 133
QY 1221 LESFIAALYTDKDLVYHVTFMNVCFFPLRK--EFILNQDNDPKSQLOQCCLTLRTEGKE 1278
Db 134 FEAFVGLYLDQGLSDSVTFABKVIFFPYVEDDELGVGVDFF---KTQFQE---YVHSQNK- 186
QY 1279 PDILPYKTLQTVGSPSHARTYTVAVYFKGERIGCGKGPSIQQAEMGAAMD 1328
Db 187 GDV-TYQLIKEGFAHRLFTSEVILEKAVAEKGKTKKESQKAAEQ 235

RESULT 3

US-11-074-176-136
; Sequence 136, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.

; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perli, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stresses-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-136

Query Match 3.0%; Score 221.5; DB 7; Length 228;
Best Local Similarity 29.0%; Pred. No. 1.3e-07;
Matches 67; Conservative 47; Mismatches 102; Indels 15; Gaps 6;
QY 1107 LTFEFAIGVIFTHVRLARAFRTV--GNHLLTGHQNMFLGDSIMOLVATEYL 1164
Db 7 LANKLNEYNIENNEKLEEAFTHSYSNEHPDDGIRDYKLEFLGDAVLELAVSNLYR 66
QY 1165 HPFDHGHHLTLRLSSLVNNRTQAKVAELGMQ--EYATINDTKRPVGLRTKTLADLLES 1223
Db 67 HYPKLNNEGTLRLSRNIVRTGFEFAIECGFQKEIHLGNBEEKAGARNRKTLLLEDVFEA 126
QY 1224 FIAALYTDKDLVYHVTFMNVCFFPLRK--EFILNQDNDPKSQLOQCCLTLRTEGKEPDI 1381
Db 127 FNGALFDQDQPAVERFLHTVYPLIAKGEF---DSDRYKTKDLQZ---LLOQNGPVNI 179
QY 1282 PLYKTLQTVGSPSHARTYTVAVYFKGERIGCGKGPSIQQAEMGAAMDALKEY 1332
Db 180 EYSVISESQLPSH---FVVELKINDQVQTQGEHNNKKAEEQQAALQKF 227

RESULT 4

US-11-059-982-1
; Sequence 1, Application US/11059982
; Publication No. US20050255507A1
; GENERAL INFORMATION:
; APPLICANT: Jenkins, Robert B.
; APPLICANT: Yang, Ping
; APPLICANT: Thibodeau, Steve
; APPLICANT: Wang, Liang
; APPLICANT: Schaid, Daniel
; TITLE OF INVENTION: CYTOGENETICALLY DETERMINED DIAGNOSIS AND
; TITLE OF INVENTION: PROGNOSIS OF PROLIFERATIVE DISORDERS
; FILE REFERENCE: 07039-505001
; CURRENT APPLICATION NUMBER: US/11/059,982
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: US 60/545,573
; PRIOR FILING DATE: 2004-02-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-059-982-1

Query Match 2.9%; Score 221; DB 7; Length 1560;
Best Local Similarity 21.7%; Pred. No. 1.6e-06;
Matches 166; Conservative 55; Mismatches 231; Indels 314; Gaps 37;
QY 4 GNTCHMSFHPGRCPRGRGGHGARGSPAFSPQNLRLHPQPPVQYQVEPPSAPST-- 61
Db 786 GDSPLPLSPHPTR-----PPSRPSPRQSVS-RPPSEPPL-HPCPPPOAPPTLP 832

QY 62 ---TFNS-----PAPNLPDRPDPVPPPPMPPSAQ---GPLPPCPPIRPPFNH----- 105
Db 833 GIFVIQNLQGVPPASNPAPATAG--PPQPLRPQSPGPPPLPPAPHPPLPPSTSSAVAS 890
QY 106 ----QMRHPPVPVPCPPMPCPNPPVPCAPPQCGTFPPMPPPPSPMPPPPPP 157
Db 891 SSETSLPAPTSDQLQPPPGGPKSTP--PTL-----HLVPEAPPPPPPTTQ 944
QY 158 ----PVMPPQ-----VNYQ 167
Db 945 MVTTTFFPALPQKALLERHQVPSGIILQKAGAPAAPOTSTSLGLTSPAASVLVSGQ 1004
QY 168 YPGY-----SHNFP-----PPSFNSQONPSSFLPSAN-----NSSSPHRHLPPYP 211
Db 1005 APSGTTAPSHAPAPMAATGLPPLPAENKAFASNLPLTNVAKAASGPGKPSGLQYE 1064
QY 212 ----LPKAPSERRSPER--LKHVDHR-----HRDHSHG-----RGERHRSDDR--BRGRS 255
Db 1065 SKLSGLKKPPTLOPSKEACFEHLHKHOGSVLHPDYKTAPPSPEDALHRLLPVHYVQAL 1124
QY 256 PRRRQDSRYRSDYDRGRTSPRHSRYSRERERHRUR-----DNRR--SPSLR----- 304
Db 1125 PS-----PSDYHKVDEEFETVSTQLLKRTQMLNKYRLLLLESRRVSPSAEMWID 1176
QY 305 -----RSYKKEYRSGRSGYGLSVVPAGCTPGLPGLIKNTDWSAP 346
Db 1177 RMFIQEBKTTALDKLAKPKDEYVSSRSLGLPTAASSEG--HRLPGH--GPLSSAP 1232
QY 347 -----PLEIVNHRSPSRKKRAREEEDKRWSDNOSCKDKNYTSIKEKEPEETM 396
Db 1233 GASTQPPPHLPKLVIRHGAGGSPVTVARASSLSSSSSSSAASLDADEGDP---M 1289
QY 397 PKNE-----EEBELLKPYW-----IRCTSHENYSSDPMQVQ----- 412
Db 1290 PSNRPPIKTYEARSIGLKLKIQEAGLSKVHVNTALDPVHPQPPPPATLKVAEPPPRP 1349
QY 413 -----IRCTSHENYSSDPMQVQ----- 431
Db 1350 PPPPTGQNGTVDPHPHPAAPERKPLGTAPHCPRLPKTYEN-----VGGPGA 1400
QY 432 -DSTVVGTSR-----LRDLVKFEELGSRQEK-----AKAAR-- 463
Db 1401 PGTAPRAGGSPAPLPKVDATSGLIREL--AAVEDELYQRLMKGPPEPAASAAQGT 1459
QY 464 --PPWBPPTKLDLEDLESSESECEDESDTSCSSSDSEVFDVIAE 507
Db 1460 GQPDWEAP-----GLPPAKRKSESPPVDVQASFSSDSPQDDTLTE 1499

RESULT 5
US-10-467-657-2638
; Sequence 2638, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2638
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2638

Query Match 2.8%; Score 211; DB 6; Length 244;
Best Local Similarity 28.3%; Pred. No. 6.7e-07;
Matches 68; Conservative 41; Mismatches 117; Indels 14; Gaps 6;
QY 1095 ROLIETSPVL--OKLTFEFAEAGVIFTHVRLLARAFTLRTVGFNHLTLGHNQMEFLGDS 1152
Db 2 RKTVMKDDVLKQQAHTAIQKLGAFRDMSLRLALTHRS--HHAK--HNEAFEPVGD 56
QY 1153 IMQLVATEYLFTHPPHGHGLTLRSSLVNNTQAKVABEELGMOEVAITNDKTRPVGL 1212
Db 57 ILNYTVARMFLFOAPPKLTGELSRLASLVNEGVLAEAAEMNVGDLGLGAGELKSSGF 116
QY 1213 RYKT-LADLLESFIAALYTDKLEYVHTFMNVCFPRLKKEFILNQDWNDDPKSOLQOCCLT 1271
Db 117 RRSIILADAEAMFAVSDADFNATKAVKVRHLFAERVRADFQNAQDKTALQEA--- 173
QY 1272 LRTEGKEPDIPLYKTLQTVGPPSHARTYTVAVYFKGERIGC-GKGPSIQQAEMGAAMDALE 1330
Db 174 --LQARRFALPKYRIEIQIGHADDSMFVISCDLGELGFVCRAGKTSRKAQAEQAEALK 231

RESULT 6
US-10-770-726-49
; Sequence 49, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-49

Query Match 2.8%; Score 207; DB 6; Length 795;
Best Local Similarity 26.4%; Pred. No. 5.3e-06;
Matches 89; Conservative 41; Mismatches 127; Indels 80; Gaps 13;
QY 214 KAPSRERSPERLKHYYD-----HRRDHS-----HGRGHRH-SLDRE 251
Db 86 KPQQMSRKKEVHHKDEKRRKRRHSAGGKHARVKEKEREKRRKHREEQDKAR 145
QY 252 RGRSPDRRQDSRYRSDYDRGRTPTSPRHSYERSR-----ERER-----ERHHR 295
Db 146 REWERQKREMAHREHSRRERDRLEQLERKREKRMREQKEQKEQKERRERRAEKERR 205
QY 296 DNRRSPSL-ERSYKKEYKRSYGLSVVPEPAGCTPGLPGLIKNTDWSAPLEIVNHR 354
Db 206 EARREYSAHRTWREDYSDVKASHWSRSP-----PRPRERFELGDKRP-----GEA 254
QY 355 SPSREKKRARWEEK-----DRWSDNOSCKDKNYTSIKEKEPEETMPDKNEEEEEELKP 410
Db 255 RPARAQKPAQLKEEKWEERDLSDLDQIDSDSRKTSABSSSSAABSSGSGSEEEEEEE 310
QY 411 VMIRCTSHENYSSYSDPMQVGVSTVGTSLRDLVDKFEELGSGRQEKAKAARPPWPPK 470
Db 311 -----EEEEEG-STSESEEEEEEEEEEEEEETGNSNEEA-SEQSAEVS 354
QY 471 TKLDEDLSSSES-----ECSEDESDTSCSSSDSEV 501
Db 355 EEMSEDEERENHLLVPESRFDRDSCGSEBAEEV 391

RESULT 7
US-10-995-561-988

```
; Sequence 988, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 988
; LENGTH: 1823
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-988

Query Match      2.5%; Score 190; DB 6; Length 1823;
Best Local Similarity 21.2%; Pred. No. 0.00018;
Matches 170; Conservative 47; Mismatches 258; Indels 328; Gaps 39;

QY 19 PRGRG-GHGARPSA-----PSFRPQNRLHLHPQPPVQYQYE-----54
Db 284 PKGDGIGPTRQPPSQGLGYPKYKSLPRFRQOQQEQLLKQQQHQHQHQQSAPPTPV 343
QY 55 PSAPS-TTFSNSPAPNFLPP-----RPDFVPPPPMPPSAQGP-----92
Db 344 PPSPPQVTLGAVPAPQAPPPPKALYPGALGRPPMPMPPMPPMPPMPPYVDPRLQ 403
QY 93 LPCCPIRP-----PPPNH-----QMRHPPVPVPCP-----117
Db 404 RPPLDYPPGVHPGSLVPRERSDSSGSSSEPPDRHAPAMLERGTTPVDPKLAWGDVFT 463
QY 118 -----PPMPPMP-----CPNN-----PPV-----132
Db 464 ATPAEPRLTSLPQAADDDKMRSETPPVPPPPYLYSGFPENGAGPPISRFPLE 523
QY 133 -PGA-----PPGGTTFPMPPPSMHP-----PPPV 159
Db 524 EPGRPLPWPFGSDEVAKIQTTPPKKEPPKEETAQLTGPEAGKPKARGVSGGQGP 583
QY 160 MPQVNYQY--PPGYSHNFPSPNSFQNNPSSFLPSANNSSPHFRLPPVPL-----212
Db 584 RESRTETRGPRGSSRRGIPPEPGA---PPRAGPIKKPPPTKVVELPPKPLEQGD 640
QY 213 -KPASERSRPERLK-----HYDHRHRDHS-----GRGE 242
Db 641 TPKPP---KPDPLKITKGLGPKETPPNGNLSAPLRLRDYSYERVGTSRGRGRGE 696
QY 243 ---RHSRLDRERGRSPDRRRQDSR---YRSD-----YDRGTTPSHRS- 280
Db 697 YFARGGRFRGTGGRGARGARSREFRSYRFRGDDGRGGTGGPNHPPAPRGRTASETRSE 756
QY 281 ---YERSRERER-----HHRDNRRSPSLERSYKKEYKRSYGLSVVPEPAGCTP 331
Db 757 GSYYEIPKRRRQSGSETSETSHESDLAPSKAFTPK-----GTUQVPLAPPPGAPP 812
QY 332 E-LPGBIKNTDSWAPLEIVNHRSPSRREKKAARW-----EEKDRWSNQSSGDKNYT 385
Db 813 SPAPARTAGGVFTPRGV-----PSR---RCGGGRPPQVPCWSPPAKSLAPK--- 861
QY 386 SIKEPEETMPDKNEEEELKPVWIRCTHSENYSYSDPMDQV-----GSTV-----435
Db 862 ---KPPTGTLPPSKEPLKEKLI-----PGPLSPVARGSGNGSGNVGMD 902
QY 436 -----VCTSLRLDLYDKFEELGSRQEKAKAARPPWPPKTKL-----DELE-- 478
Db 903 GERPRRRRHGAQQQDKPPFR---LKQRENAARGSEKPSLTLPASAPGPEEALTTV 959
QY 479 -----SSSECESEDSTSCSSSDSEVPDVAEIKR---KKAHPDRLHD 520
Db 960 TVAPAPRAAAKSPDLNQNSQDQNEWEIASESSD-----FTSERGDKEAPPVLLT 1013
```

```
QY 521 ELWYNDPGQMDNDGPLCKCSAKAR 543
Db 1014 PRAVGTGGGGGAVPGISAMSR 1036

RESULT 8
US-10-995-561-990
; Sequence 990, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 2102
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-990

Query Match      2.5%; Score 190; DB 6; Length 2102;
Best Local Similarity 21.2%; Pred. No. 0.00021;
Matches 170; Conservative 47; Mismatches 258; Indels 328; Gaps 39;

QY 19 PRGRG-GHGARPSA-----PSFRPQNRLHLHPQPPVQYQYE-----54
Db 563 PKGDGIGPTRQPPSQGLGYPKYKSLPRFRQOQQEQLLKQQQHQHQHQQSAPPTPV 622
QY 55 PSAPS-TTFSNSPAPNFLPP-----RPDFVPPPPMPPSAQGP-----92
Db 623 PPSPPQVTLGAVPAPQAPPPPKALYPGALGRPPMPMPPMPPMPPMPPYVDPRLQ 682
QY 93 LPCCPIRP-----PPPNH-----QMRHPPVPVPCP-----117
Db 683 RPPLDYPPGVHPGSLVPRERSDSSGSSSEPPDRHAPAMLERGTTPVDPKLAWGDVFT 742
QY 118 -----PPMPPMP-----CPNN-----PPV-----132
Db 743 ATPAEPRLTSLPQAADDDKMRSETPPVPPPPYLYSGFPENGAGPPISRFPLE 802
QY 133 -PGA-----PPGGTTFPMPPPSMHP-----PPPV 159
Db 803 EPGRPLPWPFGSDEVAKIQTTPPKKEPPKEETAQLTGPEAGKPKARGVSGGQGP 862
QY 160 MPQVNYQY--PPGYSHNFPSPNSFQNNPSSFLPSANNSSPHFRLPPVPL-----212
Db 863 RESRTETRGPRGSSRRGIPPEPGA---PPRAGPIKKPPPTKVVELPPKPLEQGD 919
QY 213 -KPASERSRPERLK-----HYDHRHRDHS-----GRGE 242
Db 920 TPKPP---KPDPLKITKGLGPKETPPNGNLSAPLRLRDYSYERVGTSRGRGRGE 975
QY 243 ---RHSRLDRERGRSPDRRRQDSR---YRSD-----YDRGTTPSHRS- 280
Db 976 YFARGGRFRGTGGRGARGARSREFRSYRFRGDDGRGGTGGPNHPPAPRGRTASETRSE 1035
QY 281 ---YERSRERER-----HHRDNRRSPSLERSYKKEYKRSYGLSVVPEPAGCTP 331
Db 1036 GSYYEIPKRRRQSGSETSETSHESDLAPSKAFTPK-----GTUQVPLAPPPGAPP 1091
QY 332 E-LPGBIKNTDSWAPLEIVNHRSPSRREKKAARW-----EEKDRWSNQSSGDKNYT 385
Db 1092 SPAPARTAGGVFTPRGV-----PSR---RCGGGRPPQVPCWSPPAKSLAPK--- 1140
QY 386 SIKEPEETMPDKNEEEELKPVWIRCTHSENYSYSDPMDQV-----GSTV-----435
Db 1141 ---KPPTGTLPPSKEPLKEKLI-----PGPLSPVARGSGNGSGNVGMD 1181
```

QY 436 -----VGTSLRLDLYDKPEELGSRQEKAKAARPPWPKTKL-----DELE--- 478
DB 1182 GERPRRRHGRAQODKPPFRFR---LQOERENAAARGSGKPSLTLPASAPGPEEALTTV 1238
QY 479 -----SSSECEDEDETCSSSDSEVFDVIAEIKR--KKAHPDRLHD 520
DB 1239 TVAPAPRAAAKSPDLNSQNSDOANEWEETASESSD-----FTSRRGDKKAPPPVLLT 1292
QY 521 ELWYNDPGQMDGPKCKSAKAR 543
DB 1293 PRAVGTFGGGGGGAVPGISAMSR 1315

RESULT 9

US-10-995-561-989
; Sequence 989, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 989
; LENGTH: 2108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-989

Query Match 2.5%; Score 190; DB 6; Length 2108;

Best Local Similarity 21.2%; Pred. No. 0.00022;

Matches 170; Conservative 47; Mismatches 258; Indels 328; Gaps 39;

QY 19 PRGRG-GHGARSA-----PSFRQNLRLHLPQPPVQVQYE----- 54
DB 563 PKGDGIGTPRQPPSGGLGYPKYQKSLPRFRFQOQEQQLKQOQHQOQOQSGAPPTPV 622
QY 55 PPSAPS-TTFSNSPAPNPLPP-----RPDFVPPPPMPPSAQGP----- 92
DB 623 PPSPPQPVTLGAVPAPQAPPPPKALYPCALGRPPPPMPPMPPMPPMPPVYDPRLLQG 682
QY 93 LPCCPIRP-----PPPNH-----QMRHPPVPVPCF----- 117
DB 683 RPPLDFYPPGVHPSGLVPRERSDSSGSSSEPPDRHAPAMLRERGTTPVDPKLAWGDVFT 742
QY 118 -----PPMPPMP-----CPNN-----PPV----- 132
DB 743 ATPASPRPLTSLRQAADDDKGRMSETPPVPPPPYLYSGFPENGAGPPISRFPLE 802
QY 133 -PGA-----PPQGTFFPMPPPSMHP-----PPPPV 159
DB 803 EPGPRPLMPPPGSDVAKIQTTPPKKPKBEATQLTGPEAGKPKARGVSGGQGGPPPPR 862
QY 160 MPQVNYQY--PPGYSHNFPPPSFNNSFQNNSSFLPSANNSSSHFRHLPPVPL----- 212
DB 863 RESRTETRWGPRGSGRRGIPPEEPCA---PPRAGPIKPPPPPKVBEELPKPLEQGD 919
QY 213 -PKAFSERRSRLK-----HYDDHRRDHS-----GRGE 242
DB 920 TKPP-----KPDPLKITKGLGPKETPPNGNLSAPRLRRDYSYERVGPTSCRGGRGE 975
QY 243 ---RURSLDRRGRASPPRRQDSR---YRSD-----YDRGTPSRHS-- 280
DB 976 YFARGRGFTGYGGRGARGARSREFRSYREFRGDGRGGTGGGNHPPAPRGRTASETRSE 1035
QY 281 ---YERSRERER-----HHRONRRSPSLERSYKKEYKSGRSYGLSVVPEPAGCTP 331
DB 1036 GSEYBEIPKRRQRQSGTSGTSETHESDLAPSDKEAPTPE-----GTLTQVPLAPPPGAPP 1091

QY 332 E-LPGEIINKWDSWAPPLEIVNHRSPSRREKKRABW-----EEEKDWSNQSSGDKQNT 385
DB 1092 SPAPARFTARGRVFTPRGV-----PSR---RGRGGRRPPQVPCGMSPPAKSLAPK--- 1140
QY 386 SIKEKEPETMDKNEEEELLPWIRCTHSENYSSDPMDOV-----GDSVTY----- 435
DB 1141 ----KPTGPLPSPKEPLKEKLI-----PGPLSPVARGSGSGSNVGMED 1181
QY 436 -----VGTSLRLDLYDKPEELGSRQEKAKAARPPWPKTKL-----DELE--- 478
DB 1182 GERPRRRHGRAQODKPPFRFR---LQOERENAAARGSGKPSLTLPASAPGPEEALTTV 1238
QY 479 -----SSSECEDEDETCSSSDSEVFDVIAEIKR--KKAHPDRLHD 520
DB 1239 TVAPAPRAAAKSPDLNSQNSDOANEWEETASESSD-----FTSRRGDKKAPPPVLLT 1292
QY 521 ELWYNDPGQMDGPKCKSAKAR 543
DB 1293 PRAVGTFGGGGGGAVPGISAMSR 1315

RESULT 10

US-10-995-561-991
; Sequence 991, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 991
; LENGTH: 2157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-991

Query Match 2.5%; Score 190; DB 6; Length 2157;

Best Local Similarity 21.2%; Pred. No. 0.00022;

Matches 170; Conservative 47; Mismatches 258; Indels 328; Gaps 39;

QY 19 PRGRG-GHGARSA-----PSFRQNLRLHLPQPPVQVQYE----- 54
DB 618 PKGDGIGTPRQPPSGGLGYPKYQKSLPRFRFQOQEQQLKQOQHQOQOQSGAPPTPV 677
QY 55 PPSAPS-TTFSNSPAPNPLPP-----RPDFVPPPPMPPSAQGP----- 92
DB 678 PPSPPQPVTLGAVPAPQAPPPPKALYPCALGRPPPPMPPMPPMPPMPPVYDPRLLQG 737
QY 93 LPCCPIRP-----PPPNH-----QMRHPPVPVPCF----- 117
DB 738 RPPLDFYPPGVHPSGLVPRERSDSSGSSSEPPDRHAPAMLRERGTTPVDPKLAWGDVFT 797
QY 118 -----PPMPPMP-----CPNN-----PPV----- 132
DB 798 ATPASPRPLTSLRQAADDDKGRMSETPPVPPPPYLYSGFPENGAGPPISRFPLE 857
QY 133 -PGA-----PPQGTFFPMPPPSMHP-----PPPPV 159
DB 858 EPGPRPLMPPPGSDVAKIQTTPPKKPKBEATQLTGPEAGKPKARGVSGGQGGPPPPR 917
QY 160 MPQVNYQY--PPGYSHNFPPPSFNNSFQNNSSFLPSANNSSSHFRHLPPVPL----- 212
DB 918 RESRTETRWGPRGSGRRGIPPEEPCA---PPRAGPIKPPPPPKVBEELPKPLEQGD 974
QY 213 -PKAFSERRSRLK-----HYDDHRRDHS-----GRGE 242
DB 975 TKPP-----KPDPLKITKGLGPKETPPNGNLSAPRLRRDYSYERVGPTSCRGGRGE 1030

QY 243 ---RHSLDRRGRSPDRRRQDSR-----YRSD-----YDRGRTSPHRHS- 280
Db 1031 YFARGRGFGTYGGRGARGSRFRSRYRFRGDDGRGGTGGNHPAPRGRGTATSE 1090
QY 281 ---YERSRERER-----HRHRNRRSPSLERSYKKEYKRSRSGLSVVPVAGCTP 331
Db 1091 GSEYBEIPKRRQORGETSETHESDLAPDKEAPTPKE---GTLTQVPLAPPFGAPP 1146
QY 332 E-LPGSEIIKNTDWAAPLBIVNHRSRSPREKKAARW-----EEXKRWSONSGKDKNXT 385
Db 1147 SPAPARFTARGRVFTPGV-----PSR---RGRGGGRPPQVCPGWSPPAKSLAPK--- 1195
QY 386 SIKEKPEBTMDKNEEBEELKPVWIRCTHSENYSSDPMDOV-----GDSTV---- 435
Db 1196 ----KPTGFLPPSKPELKEKLI-----PCPLSPVARGSGNGSGNVGMD 1236
QY 436 -----VGTSLRDLYDKFEEELGSRQEKAKAAPPPWEPKTKL-----DEDL- 478
Db 1237 GERPRRRHGRAQQQDKPPFR--LQERENAAARGSEKSLTLPASAPGPEALITV 1293
QY 479 -----SSSECEDEDESDSCSSSDSEVDFVIAEIKR--KXAHPPRLHD 520
Db 1294 TVAPAPRAAAKSPDLNSQNSQANEWEETASESSD-----FTSRRGDKEAPPVLLT 1347
QY 521 ELWYNDPGQWNDGFLCKCSAKAR 543
Db 1348 PKAVGTPGGGGGAVPGISAMSR 1370

RESULT 11

US-10-821-234-1074
; Sequence 1074, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1074
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1074

Query Match 2.3%; Score 175.5; DB 6; Length 175;
Best Local Similarity 29.2%; Pred. No. 7.9e-05;
Matches 57; Conservative 13; Mismatches 52; Indels 73; Gaps 12;
QY 62 TFSNSPAPNPLPRPDPVFPFPPMP-PSAQGLPPCPIRPPFPHQMRHPPFVPPCFPPM 120
Db 24 THLAMNPNAGQPGPN--PYPPNIGCGGSPNAPHPPIPNPF----- 63
QY 121 PPMPCNNPPVPGAPPGQGTFFPMMPSPSMHPHPPPPVMPQ-QVNYQVPPGYSHNFPP 179
Db 64 -PPGPC---PPPGAPHGNAF----PPGGPHFVPPQGYPGCQPIGPVPPPY-----PP 110
QY 180 PSFNSFNQNSFLPSANNSSPHRHLPLPYLPKAPS-----ERRSPRLK----- 226
Db 111 PA-----PG--IPPVN-----PLAPGVGPAVIVDKMKMKKAKKK 146
QY 227 -HYDDHRRDHSRGR 240
Db 147 MHKHQKHKKYKHGK 161

RESULT 12
US-10-821-234-1641
; Sequence 1641, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1641
; LENGTH: 1970
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1641

Query Match 2.3%; Score 174; DB 6; Length 1970;
Best Local Similarity 27.8%; Pred. No. 0.0021;
Matches 64; Conservative 27; Mismatches 101; Indels 38; Gaps 13;
QY 19 PRGRGHGARGAPSRFRPQNLRLHLPQQPPVQYQYEP-----PSAPS---TTFNSNP-A 68
Db 1602 PRSPGY--TPQSPSYSPS-----PSYSPTSPSYSPSPTSPSYSPSPTSPSYSPS 1654
QY 69 PNFLLPRDPFVFPFPPMPSPSAQGLPPCP-IRPPFNHQMRRHP--FPVPPCFFPMPP--- 122
Db 1655 PSYSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPS 1714
QY 123 ---PMPCPNNPPV-PCAPPGQGTFFPMMPSPSMHPHPPPPVMPQVNYQ-YPPGY--HH 175
Db 1715 PTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSP 1774
QY 176 NFPFSPFNSFQNNPS-----SFLPS--ANNSSPHFRHLPPVPLPKARS 217
Db 1775 NYTPTSPNYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPSPTSPSYSPS 1824

RESULT 13

US-10-821-234-914
; Sequence 914, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 914
; LENGTH: 1532
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-914

Query Match 2.3%; Score 173; DB 6; Length 1532;
Best Local Similarity 28.3%; Pred. No. 0.0017;
Matches 82; Conservative 15; Mismatches 97; Indels 96; Gaps 18;
QY 13 HPGRCGPRGRGG-----HGAR-PSAPSRFPQNLRLHLPQQPPVQYQYEP-----P 56

This Page Blank (usp1c)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 00:51:00 ; Search time 52 Seconds
(without alignments)
2184.546 Million cell updates/sec

Title: US-10-774-974-2

Perfect score: 7500

Sequence: 1 MQQNTCHRMSPHGRGCRP.....MRWREHQREPDTEYDIKK 1374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5 COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/6 COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/H COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/RE COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7500	100.0	1374	2	US-09-900-4258-2
2	2463	32.8	466	2	US-09-900-4258-37
3	890	11.9	412	2	US-09-900-4258-3
4	496	6.6	115	2	US-09-513-599C-4635
5	316	4.2	553	2	US-09-949-016-7961
6	312.5	4.2	241	2	US-09-107-532A-5434
7	284.5	3.8	232	2	US-09-583-110-4554
8	284.5	3.8	236	2	US-09-107-433-4394
9	284.5	3.8	998	2	US-09-949-016-7757
10	277.5	3.7	232	1	US-08-869-674-2
11	277.5	3.7	232	2	US-09-213-010-2
12	277.5	3.7	232	2	US-09-213-011-2
13	271.5	3.6	386	2	US-09-461-774-10
14	267	3.6	243	2	US-08-965-087-2
15	264.5	3.5	1404	2	US-08-801-308-1
16	263	3.5	227	2	US-09-489-039A-11549
17	262	3.5	226	2	US-09-900-4258-6
18	262	3.5	971	2	US-09-248-796A-19531
19	260.5	3.5	245	2	US-09-710-279-1658
20	260.5	3.5	246	2	US-09-134-001C-4787
21	254.5	3.4	607	2	US-09-248-796A-19157
22	252	3.4	228	2	US-09-543-681A-7646
23	248.5	3.3	904	2	US-09-376-594-615
24	239.5	3.2	439	2	US-10-300-819B-21
25	239.5	3.2	865	2	US-09-281-766-19
26	239.5	3.2	865	2	US-09-612-858-19
27	239.5	3.2	865	2	US-09-957-995A-19

28	239	3.2	253	2	US-09-352-991A-29495	Sequence 29495, A
29	233.5	3.1	288	2	US-09-340-236-2003	Sequence 2003, Ap
30	232.5	3.1	1646	2	US-09-535-008-67	Sequence 67, Appl
31	232.5	3.1	1647	2	US-09-535-008-2	Sequence 2, Appl
32	232.5	3.1	1647	2	US-09-824-574-4	Sequence 4, Appl
33	232.5	3.1	1647	2	US-09-538-092-1172	Sequence 1172, Ap
34	232.5	3.1	1649	2	US-09-535-008-75	Sequence 75, Appl
35	232.5	3.1	1650	2	US-09-535-008-71	Sequence 71, Appl
36	232.5	3.1	1659	2	US-09-949-016-9752	Sequence 9752, Ap
37	232.5	3.1	1678	2	US-09-535-008-69	Sequence 69, Appl
38	232.5	3.1	1679	2	US-09-535-008-65	Sequence 65, Appl
39	232.5	3.1	1681	2	US-09-535-008-77	Sequence 77, Appl
40	232.5	3.1	1682	2	US-09-535-008-73	Sequence 73, Appl
41	230.5	3.1	254	2	US-09-553-863-18	Sequence 18, Appl
42	230.5	3.1	1909	2	US-09-590-968B-2	Sequence 2, Appl
43	230	3.1	666	2	US-09-050-739-70	Sequence 70, Appl
44	229.5	3.1	424	2	US-09-538-092-1338	Sequence 1338, Ap
45	228.5	3.0	259	2	US-09-438-185A-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1

US-09-900-4258-2
; Sequence 2, Application US/09900425B
; Patent No. 6737512
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425B
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-4258-2

Query Match		100.0%;	Score 7500;	DB 2;	Length 1374;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1374;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MQQNTCHRMSPHGRGCRGCGHGRGARGP	SAPSRPQNLRLHLPQPPVQYQYEPSPAPS	60	
Db	1	MQQNTCHRMSPHGRGCRGCGHGRGARGP	SAPSRPQNLRLHLPQPPVQYQYEPSPAPS	60	
QY	61	TTFSNSPAPNLFPPRDPFVPPPPMPPSAQGLP	PCPIRPPFPNQHMRHPPVPPCFPPM	120	
Db	61	TTFSNSPAPNLFPPRDPFVPPPPMPPSAQGLP	PCPIRPPFPNQHMRHPPVPPCFPPM	120	
QY	121	PPMPCPNPPVPGAPGCGTFFMPPSPHPPPP	PPMPPVMPQOVNTQYPGYSHNPPPP	180	
Db	121	PPMPCPNPPVPGAPGCGTFFMPPSPHPPPP	PPMPPVMPQOVNTQYPGYSHNPPPP	180	
QY	181	SFNSFQNNSSFLPSANNSSPHFHLPPYPLP	KAPSERRSERLKHYYDDHRRDHSRGR	240	
Db	181	SFNSFQNNSSFLPSANNSSPHFHLPPYPLP	KAPSERRSERLKHYYDDHRRDHSRGR	240	
QY	241	GERHSLDRRGRGPDRRRQDSRYSDYDRGRT	PSRHSYERSRERERHHRDRNR	300	
Db	241	GERHSLDRRGRGPDRRRQDSRYSDYDRGRT	PSRHSYERSRERERHHRDRNR	300	
QY	301	PSLEHSYKKEYKRGSGSYGLSVVPPAGCTP	ELPGELIKNTDSWAPPLIVNHRSPSREK	360	
Db	301	PSLEHSYKKEYKRGSGSYGLSVVPPAGCTP	ELPGELIKNTDSWAPPLIVNHRSPSREK	360	
QY	361	KEARWEEKDRWSDNQSSGDKNYTSIKKEP	ETMPDKNBEELLLKPVWIRCTHSEN	420	
Db	361	KEARWEEKDRWSDNQSSGDKNYTSIKKEP	ETMPDKNBEELLLKPVWIRCTHSEN	420	


```
Db 2 SLFNIWKTSGGEP-----ILNRLRYLGDVAVVELIVSHLLYFNLTHFEGGLATRYAL 57
QY 1004 VQOHLAMLAKELEDPFMYLAHGPDLCRSDLRHAWANCFEALICAVYLEGSLEBAKOL 1063
Db 58 VQNRNLATLAKNCRIDEMLOYSHGADLINVABFKHALANAFEAVMAAIVLDGGLAPCDVI 117
QY 1064 FGELLP-NBDPLREVLNPLPLQLOEPNTDQLIETSPVLOKLTFFEAATVIFTHVR 1122
Db 118 FSKAMYGHQVLEKWDHNEHEKREDQGDRLDSFITFTLSTFHALERLGIQNNIR 177
QY 1123 LLARAPLTRVGNHLLTLGHNORMEFLGDSIMOLVATEYLFTHFPDHHGHLTLRSSLIV 1182
Db 178 LLAKATRNIPNNDLTGHNQLEWLGSLVLIQVDSFLYRFPYTHHGHWSLLTSIV 237
QY 1183 NNRTQAKVABELGMOYAIITNDKTRPV---GLRTKTLADLLESFIAALYTDKDLVYHT 1239
Db 238 SNOTQAVVCDLGLTFEVI-----KAPYKTPPELKLKADLVEAFIALVVDRIEHCRA 292
QY 1240 FMVVCFFPRLKEFILNDNDPKSQOCCLTLR-TEGKEPDILYKTLQTVGSPHARTY 1298
Db 293 FIRIVFCPLKHFIESEKWDKSHLQWCLAMRDPSSSEPDMPYRVLGIEGPTNNRIF 352
QY 1299 TVAVYKGRIRGGKGPSIQOAEWGA--AMDALRYNFPOM-AHQKRFGRKYROELKE 1354
Db 353 KIAVYTKGRILASAESNVHKAELRVAEALANLESMSFKKAKNNSWRRLEQDTS 412
```

```
RESULT 4
US-09-513-999C-4635
; Sequence 4635, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
```

```
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4635
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23...-1
; OTHER INFORMATION: score 5.9
; OTHER INFORMATION: seq TLKFLTLQKSN/KR
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 3
; OTHER INFORMATION: Xaa=Ala or Pro
US-09-513-999C-4635
```

```
Query Match 6.6%; Score 496; DB 2; Length 115;
Best Local Similarity 95.7%; Pred. No. 2.2e-31;
Matches 88; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 511 KKAHPDLRDLWYNDPQNDGPKCKSAKARRTGIRHSIYPGEAIPKPCPMTNACR 570
Db 24 KRXNLDRLHDELWYNDPQNDGPKCKSAKARRTGIRHSIYPGEAIPKPCPMTNACR 83
QY 571 LFHYRITVSPPTNFLTDRPTVIEYDDHEVIFE 602
Db 84 LFHYRITVSPPTNFLTDRPTVIEYDDHEVIFE 115
```

```
RESULT 5
US-09-949-016-7961
; Sequence 7961, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7961
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7961
```

```
Query Match 4.2%; Score 316; DB 2; Length 553;
Best Local Similarity 28.0%; Pred. No. 3.2e-16;
Matches 111; Conservative 32; Mismatches 106; Indels 148; Gaps 21;

QY 19 PRGRGHGARGAPSRFPQNTLLHPQPPVQVQYPPSPASTTFSNSPAPNF---LPP 74
Db 200 POGGRGRGFGA-----VPGDRFPGPAGP-----GGPPPPFPAGQTPP 239
QY 75 RDPFVPPPPMPSPSAGPLPPCIRPP-----PFIHQMRHPPV-----PPC--FPP 119
Db 240 RP---PLGPPGPPGPPPPGQVLPPLLAGPFRNGDRPPVLPVLPQGFQPPGLGLPP 296
QY 120 MPPMPCPNNPVPGAPCQ-----GTFP-----FMPPPPSPHPPPP--PPV 159
Db 297 GPPP-FVPGYGPVPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 355
QY 160 MPQQVNYQVPGYSHNPPSPFSQNNPSPFLSANSSSP---HFRHLPPY----- 210
Db 356 AP-HVN-----PAF-----FPPPT-----NSGMPTSDSRGPPPTDPPYGRPPYDRGDYG 398
QY 211 PLPKAPSRSPERLKHYYDHRDHS-----HG-RGERHRSUDRRERGRSPRRRQDSRYRSDYDRGTP 458
Db 399 PPGREMDTARTPLSEAEPEEIMNRNRAISSAISRAVSDASAGDYSAJETLVTAISLIK 275
QY 238 -----HG-RGERHRSUDRRERGRSPRRRQDSRYRSDYDRGTP 275
Db 459 QSKVSADDRCKVLISLQDLHGIESKSYGSGSRERSRERDHSR--SREKSRHKSRSR 516
QY 276 SEHRSYERSRERERHRRDNRSPSLERSYKKEYK 312
Db 517 DRHDDYRERSRERHRRDRD--RDRDRDREREREYR 551
```

```
RESULT 6
US-09-107-532A-5434
; Sequence 5434, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
```



```
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...236
; SEQUENCE DESCRIPTION: SEQ ID NO: 4394:
US-09-107-433-4394

Query Match
Best Local Similarity 32.9%; Pred. No. 2.8e-14; Length 236;
Matches 76; Conservative 48; Mismatches 90; Indels 17; Gaps 8;

Qy 1108 TEPEEAGVLPVTHVLLARAFATLTTCVGNH--LTIGHNORWELGDSIMOLVATEYLPIH 1165
Db 10 TVLNKHPAIEFTDKLLETAPTHTSYANEHRLKISHNERLEFLGDVAVQLLSELYKK 69

Qy 1166 FPDHHEGLTLRLSSLVNNTQAKVAEELGQVEYAITNDKTRPVGLRTK--TLADLLES 1223
Db 70 YPKPEGDLKRAMIVREESLAGFARDQDFP-KLKGKEKSGGRENRTILGDAFEA 128

Qy 1224 FIAALYTDKOLEVYHTPMNVCFPPRLK--BFILNQMDNDPKSOLQCCCLTLRTGKEPDI 1281
Db 129 FLGALLLDKDAVKKEPIYQWMPKVEAGEFEMITDY---KTHLQE---LLQVNG---DV 179

Qy 1282 PL-YKTLQTVGPSHARYTYAVYFKGERICGCKGPSIQQAEMGAMDALEK 1331
Db 180 AIRYQVISETPAHDKVFDEVLVEGKSIGQGGRSKGLAEQAAKNAVEK 230

RESULT 9
US-09-949-016-7757
; Sequence 7757, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7757
; LENGTH: 998
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7757

Query Match
Best Local Similarity 22.7%; Pred. No. 2.3e-13; Length 998;
Matches 151; Conservative 63; Mismatches 206; Indels 245; Gaps 30;

Qy 27 ARPSA--PSRPNQLRLHQQBPVQVQYEPSPA-----PSTTFNSPAPNELPPR 75
Db 350 ARPGGGRPGHEHN-KLGYLVSPQQIRRGERSCYRSINGRHHSERSQRTQGPS-LPAT 407

Qy 76 PDFVPPPPPPPSAQQGLPCPIRPPPPNQHRRHPVPPPCFPPMPMPMPCNNPPVPGA 135
Db 408 PVFVPVPPP-----PLYPPPP-----HTLPLP---PCVPPQPSQPP--PGQ 445

Qy 136 PPGGTGTFPPMPPMPPMPPMPPMPPMPPMPPMPPMPPMPPMPPMPPMPPMPPMPPMPP 170
Db 446 PPPAG---YVPPPPGFP-PAPANLSTWSSGVQTAHSNTIPTTQAPLSEEFYREQR 501

Qy 171 -----GYSHNPPPPSPNS 184
Db 502 LKBEKKKSKLDBFTNDFAKELMEYKIKERRRSFSRSKSPYSGSSYSRSYTSKRS 561
```

```
Qy 185 FQNPSSFLPSANNSSPHRHLPPYPLPKAPSRSPERLKHYYDHRHDHSHGERH 244
Db 562 GSTRSRYSRFSRSHSRYSRSPYP-----RRGRGKSRNY---RSRSRSHG---YH 608
Qy 245 RSLDRRGRSPDRRQDSRYSDYD-RGRTPSRHSYRSRER-----289
Db 609 RS-----RSRSPYRYHSRSPQAFRGQSPKRVNVPQGETREYFNRYREVPYDMK 663
Qy 290 -----ERHRDNRRSPSLERSYKKEYKRSRSGYGLSVVPEPAGCTPELPG 336
Db 664 AVYGRSDVPDPPEKERYE-----WERKYREWEYKYGYAAGAPRPSA-----709
Qy 337 IINKTUSWAP-----PLEIVNHRSPSREKKARWEEKRWSDNQSSG-----KD 381
Db 710 ---NRENFSPERFLPLNIRN--SPFTRGRREDTVGGQSHRSRNYGNYPEKLSARDHNQ 764
Qy 382 KNTYSIKEPESTMPD-----KNEEEEBELLKPVWIRCTH-----SEN 421
Db 765 KONTKEKESENAPOGDKGNKHKKRRKRGSEGLNPELLETSRKSRPTGVEENK 824
Qy 422 YSS-----DPMQVGDSTVVGTSLRLDLVDKFEEL-----GSRQ 456
Db 825 TDSLFLVPSRDDATPVDRDPM-----AESITFKSVSEKDKRERDKPAKGDKTKRKNDSAV 882
Qy 457 EKAKAARPWPPEPKTKLDSLESSESEC-----ESDEDSTCSSSSDSEVEDVIAEK 509
Db 883 SKKENTVTKPAKGQEKVDGERERSRSEPPKIKAKEETPKTNDTKSSSSQDEKITGTP 942

510 RKKAH 514
943 R-KAH 946

RESULT 10
US-08-869-674-2
; Sequence 2, Application US/08869674
; Patent No. 5866365
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/869,674
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: GM10013
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
```

QY	1108	TEFEAIGVIFTHVRLARAFTLRTVGFNH--LTLGHNQRMFGLGDSIMOLVATEYLFH	1165
DB	6	TVLKNHFAIEFADKNLLETAFTHTSYANEHRLKISHNERLEFLGDAVLQLLISEVLYKK	65
QY	1166	FPDHHEGHLTLRSSLVNNRTQAKVABELGMOEYAITNDKTRPVGLRTK--TLADLLES	1223
DB	66	YPKPEGDLKSLRAMIVREESLAGFARDQDF--IKLGGEKSGGRNDRDTILGDAFEA	124
QY	1224	FIAALYTDKOLEYVHTFMNVCFPPRLK--EFILNQDNPKSQLOQCCLTLRTGEGKEDI	1281
DB	125	FLGALLDKDVAKEFIYQVMPKVEAGEFEMITDY---KTHLQE---LLQVNG---DV	175
QY	1282	PL-YKTLOTGVPSSHARTYTVAVYFKGERIGCGKGPSIQQAEMGAAMDALK	1331
DB	176	AIRYQVISTGPAHDKVDVEVLVEGKSIGGGGRSKKLAEQAQNAVEK	226
RESULT 12			
US-09-213-011-2			
; Sequence 2, Application US/09213010			
; Patent No. 6444208			
; GENERAL INFORMATION:			
; APPLICANT: Lonetto, Michael A.			
; APPLICANT: Rosenberg, Martin			
; TITLE OF INVENTION: NOVEL COMPOUNDS			
; NUMBER OF SEQUENCES: 4			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Dechert Price & Rhoads			
; STREET: 997 Lenox Drive, Building 3, Suite 210			
; CITY: Lawrenceville			
; STATE: NJ			
; COUNTRY: USA			
; ZIP: 08543			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: DOS			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/213,011			
; FILING DATE:			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/869,674			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Bloom, Allen			
; REGISTRATION NUMBER: 29,135			
; REFERENCE/DOCKET NUMBER: GM10013			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 609-520-3214			
; TELEFAX: 609-520-3259			
; TELEX:			
; INFORMATION FOR SEQ ID NO: 2:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 232 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-09-213-011-2			
Query Match 3.7%; Score 277.5; DB 2; Length 232;			
Best Local Similarity 32.5%; Pred. No. 9.8e-14;			
Matches 75; Conservative 47; Mismatches 92; Indels 17; Gaps 8;			
QY	1108	TEFEAIGVIFTHVRLARAFTLRTVGFNH--LTLGHNQRMFGLGDSIMOLVATEYLFH	1165
DB	6	TVLKNHFAIEFADKNLLETAFTHTSYANEHRLKISHNERLEFLGDAVLQLLISEVLYKK	65
QY	1166	FPDHHEGHLTLRSSLVNNRTQAKVABELGMOEYAITNDKTRPVGLRTK--TLADLLES	1223
DB	66	YPKPEGDLKSLRAMIVREESLAGFARDQDF--IKLGGEKSGGRNDRDTILGDAFEA	124
QY	1224	FIAALYTDKOLEYVHTFMNVCFPPRLK--EFILNQDNPKSQLOQCCLTLRTGEGKEDI	1281
DB	125	FLGALLDKDVAKEFIYQVMPKVEAGEFEMITDY---KTHLQE---LLQVNG---DV	175
QY	1282	PL-YKTLOTGVPSSHARTYTVAVYFKGERIGCGKGPSIQQAEMGAAMDALK	1331
DB	176	AIRYQVISTGPAHDKVDVEVLVEGKSIGGGGRSKKLAEQAQNAVEK	226
RESULT 11			
US-09-213-010-2			
; Sequence 2, Application US/09213010			
; Patent No. 6251630			
; GENERAL INFORMATION:			
; APPLICANT: Lonetto, Michael A.			
; APPLICANT: Rosenberg, Martin			
; TITLE OF INVENTION: NOVEL COMPOUNDS			
; NUMBER OF SEQUENCES: 4			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Dechert Price & Rhoads			
; STREET: 997 Lenox Drive, Building 3, Suite 210			
; CITY: Lawrenceville			
; STATE: NJ			
; COUNTRY: USA			
; ZIP: 08543			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: DOS			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/213,010			
; FILING DATE:			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/869,674			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Bloom, Allen			
; REGISTRATION NUMBER: 29,135			
; REFERENCE/DOCKET NUMBER: GM10013			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 609-520-3214			
; TELEFAX: 609-520-3259			
; TELEX:			
; INFORMATION FOR SEQ ID NO: 2:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 232 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-09-213-010-2			
Query Match 3.7%; Score 277.5; DB 2; Length 232;			
Best Local Similarity 32.5%; Pred. No. 9.8e-14;			
Matches 75; Conservative 47; Mismatches 92; Indels 17; Gaps 8;			

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 00:51:20 ; Search time 190 seconds
(without alignments)
3021.566 Million cell updates/sec

Title: US-10-774-974-2
Perfect score: 7500
Sequence: 1 MMQGNCHRMSPHPCGCRP.....MRWERHQRDEPTEYDIK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7500	100.0	1374	3	US-09-900-425A-2
2	7500	100.0	1374	4	US-10-079-185-2
3	7500	100.0	1374	4	US-10-774-974-2
4	7500	100.0	1374	4	US-10-805-919-2
5	7500	100.0	1374	6	US-11-001-993-2
6	7486	99.8	1374	5	US-10-723-860-281
7	7486	99.8	1374	5	US-10-756-149-4809
8	2963	39.5	1327	6	US-11-097-143-2409
9	2847	38.0	541	4	US-10-205-331-116
10	2463	32.8	466	4	US-10-774-974-37
11	1961	26.1	378	4	US-10-103-313-471
12	1320	17.6	263	4	US-10-103-313-307
13	890	11.9	412	3	US-09-900-425A-3
14	890	11.9	412	3	US-10-079-185-3
15	890	11.9	412	4	US-10-774-974-3
16	890	11.9	412	4	US-10-805-919-3
17	890	11.9	412	6	US-11-001-993-3
18	340	4.5	1151	6	US-11-097-143-11586
19	311	4.1	560	5	US-10-450-763-51399
20	309.5	4.1	1557	4	US-10-437-963-137427
21	303.5	4.0	406	5	US-10-450-763-57609
22	298.5	4.0	237	4	US-10-437-963-130873
23	297.5	4.0	800	4	US-10-029-386-32198
24	292	3.9	466	4	US-10-437-963-167964
25	287.5	3.8	229	4	US-10-424-599-207316
26	285.5	3.8	258	4	US-10-437-963-159481
27	284.5	3.8	236	5	US-10-617-320-4394

28	284.5	3.8	1616	5	US-10-934-998-88	Sequence 88, Appl
29	282.5	3.8	445	4	US-10-437-963-161789	Sequence 161789,
30	282	3.8	329	4	US-10-437-963-197333	Sequence 197333,
31	281	3.7	598	4	US-10-282-122A-64954	Sequence 64954, A
32	280	3.7	266	4	US-10-425-115-189501	Sequence 189501,
33	279.5	3.7	232	5	US-10-472-928-2508	Sequence 2508, Ap
34	278	3.7	180	4	US-10-425-115-229737	Sequence 229737,
35	277.5	3.7	285	4	US-10-437-963-189482	Sequence 189482,
36	277	3.7	1537	6	US-11-097-143-27759	Sequence 27759, A
37	276.5	3.7	187	4	US-10-425-115-352789	Sequence 352789,
38	276	3.7	273	4	US-10-425-115-330708	Sequence 330708,
39	276	3.7	948	4	US-10-308-448-15	Sequence 15, Appl
40	276	3.7	948	5	US-10-934-998-295	Sequence 295, App
41	275.5	3.7	521	4	US-10-437-963-179715	Sequence 179715, A
42	275.5	3.7	594	6	US-11-097-143-10878	Sequence 10878, A
43	274.5	3.7	399	4	US-10-437-963-124007	Sequence 124007,
44	274	3.7	280	4	US-10-425-115-231956	Sequence 231956,
45	273.5	3.6	431	4	US-10-437-963-204963	Sequence 204963,

ALIGNMENTS

RESULT 1

US-09-900-425A-2
; Sequence 2, Application US/09900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-425A-2

Query Match	100.0%;	Score	7500;	DB	3;	Length	1374;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1374;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MMQGNCHRMSPHPCGCRPGRGCHGARPSA	FRPNLRLHLPQPPVQYQYPPSPAPS	60			
Db	1	MMQGNCHRMSPHPCGCRPGRGCHGARPSA	FRPNLRLHLPQPPVQYQYPPSPAPS	60			
Qy	61	TTFSNSPAPNLPDPDFVPPFPPMPSPSA	QGLPPLPCPIRPPFNHQHHPFPVPPCFPPM	120			
Db	61	TTFSNSPAPNLPDPDFVPPFPPMPSPSA	QGLPPLPCPIRPPFNHQHHPFPVPPCFPPM	120			
Qy	121	PPMPCCNPPVPGAPPGQGTFFMMPPSP	MPHPPPPVMPQVQYPPGYSHNPPPP	180			
Db	121	PPMPCCNPPVPGAPPGQGTFFMMPPSP	MPHPPPPVMPQVQYPPGYSHNPPPP	180			
Qy	181	SFNSFONNPSSTPSANNSSPHFRHL	PPVPLPKASRRSPERLKHDDHRRDHSGR	240			
Db	181	SFNSFONNPSSTPSANNSSPHFRHL	PPVPLPKASRRSPERLKHDDHRRDHSGR	240			
Qy	241	GERHRLDRRGRSPDRRQDSRYSDYDRG	RTPTSRHSYERSRERERHHRDNRRS	300			
Db	241	GERHRLDRRGRSPDRRQDSRYSDYDRG	RTPTSRHSYERSRERERHHRDNRRS	300			
Qy	301	PSLSRYKKEYKRSKSYGLSVVPEPAG	CTPELPGIILKNTDSWAPPLEIVNHRSP	360			


```
Db 301 PSLSRYKKEYKSGSYGLSVVPEAGCTPELPGHIIKNTDSWAPPLIIVNHRSPSREK 360
QY 361 KQARWEEKDRWSDNQGSGKDKNYTSIKKEPPEETMPDKNEEBEELLKPVMIRCTHSEN 420
Db 361 KQARWEEKDRWSDNQGSGKDKNYTSIKKEPPEETMPDKNEEBEELLKPVMIRCTHSEN 420
QY 421 YYSDDPMDQVGDSTVVGTSLRLDLYDKFEEELGSRQEKAKAAPPPPEPKTKLDELESS 480
Db 421 YYSDDPMDQVGDSTVVGTSLRLDLYDKFEEELGSRQEKAKAAPPPPEPKTKLDELESS 480
QY 481 SSESCESEDSTCSSSDSEVFDVIAEIKRKAHPDRLHDELWYNDPGQWNGDGLCKCSA 540
Db 481 SSESCESEDSTCSSSDSEVFDVIAEIKRKAHPDRLHDELWYNDPGQWNGDGLCKCSA 540
QY 541 KARTGIRHSIYPGEBAIKPCRMTNNAAGRLFHYRITVSPPTNFLTDRPTVIYDDHEIYI 600
Db 541 KARTGIRHSIYPGEBAIKPCRMTNNAAGRLFHYRITVSPPTNFLTDRPTVIYDDHEIYI 600
QY 601 FEGFSMFAHAPLTNIPLCKVIRENIDYTHFIEEMMPENFCVKGLELSFLFRDILELY 660
Db 601 FEGFSMFAHAPLTNIPLCKVIRENIDYTHFIEEMMPENFCVKGLELSFLFRDILELY 660
QY 661 DNLKGLPDSPPCCPRFHFMRFRVFLPDGCKEVLMSHQIILLYLLRCSKALVPEEIIA 720
Db 661 DNLKGLPDSPPCCPRFHFMRFRVFLPDGCKEVLMSHQIILLYLLRCSKALVPEEIIA 720
QY 721 NMLQWEELEWQKYAECKGMIVTNPGTKSSVRIDQLDREQFNPDPVITPPIIVHFGIRPA 780
Db 721 NMLQWEELEWQKYAECKGMIVTNPGTKSSVRIDQLDREQFNPDPVITPPIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLWKSYYVKLRLHLLANSFKVQKTDKQKLAQREBALOKIKQKNTMRREVTV 840
Db 781 QLSYAGDPQYQKLWKSYYVKLRLHLLANSFKVQKTDKQKLAQREBALOKIKQKNTMRREVTV 840
QY 841 ELSQSQCFWKTGIRSDVCOAHMMLPVLTHIRYHQCLMHLDKLIGYTFQDRCLLQLANTHP 900
Db 841 ELSQSQCFWKTGIRSDVCOAHMMLPVLTHIRYHQCLMHLDKLIGYTFQDRCLLQLANTHP 900
QY 901 SHLNFQGNPDHARNLSNGIRQPKYQKDRKVHMMMRKKGINTLINMSRLGQDDPTPS 960
Db 901 SHLNFQGNPDHARNLSNGIRQPKYQKDRKVHMMMRKKGINTLINMSRLGQDDPTPS 960
QY 961 RINHNERLEFLGDAVVEFLTSVHLYYLFPSSLEGGIATYRTAIVQNOHLAMAKKLELDP 1020
Db 961 RINHNERLEFLGDAVVEFLTSVHLYYLFPSSLEGGIATYRTAIVQNOHLAMAKKLELDP 1020
QY 1021 FMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLEAEAKOLFGRLLFNDPDLREVWLN 1080
Db 1021 FMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLEAEAKOLFGRLLFNDPDLREVWLN 1080
QY 1081 YPLHPLQLOBPNTDROLIETSPVLQKLTETEPERAGVIFTHVRLARAFTLRTVGFNHLTL 1140
Db 1081 YPLHPLQLOBPNTDROLIETSPVLQKLTETEPERAGVIFTHVRLARAFTLRTVGFNHLTL 1140
QY 1141 GHNQRMFLGDSIMQLVATEYLTHFPDHHHEGHLTLIRSSLVNVRTQAKVAEELGQOYA 1200
Db 1141 GHNQRMFLGDSIMQLVATEYLTHFPDHHHEGHLTLIRSSLVNVRTQAKVAEELGQOYA 1200
QY 1201 ITNDKTRPVGLTKTLADLLESFIAALYTDKLEYVHTFMNVCFPPRLKEFLINQDWND 1260
Db 1201 ITNDKTRPVGLTKTLADLLESFIAALYTDKLEYVHTFMNVCFPPRLKEFLINQDWND 1260
QY 1261 PKSQLQCCCLTLRTEGKEPDIPLYKTLQTVGPHSHARTYTVAVYFKGIRICGCGKPSIQQA 1320
Db 1261 PKSQLQCCCLTLRTEGKEPDIPLYKTLQTVGPHSHARTYTVAVYFKGIRICGCGKPSIQQA 1320
QY 1321 EMGAAMDALKYFPQAHQKRFIGRKYQRELKEMRWERHQBREPDETDIKK 1374
Db 1321 EMGAAMDALKYFPQAHQKRFIGRKYQRELKEMRWERHQBREPDETDIKK 1374
```

```
US-10-079-185-2
; Sequence 2, Application US/10079185
; Publication No. US2003004941A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS5030
; CURRENT APPLICATION NUMBER: US/10/079,185
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-079-185-2
```

Query Match 100.0%; Score 7500; DB 4; Length 1374;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MMQGNCTCHMSFHPGRGCPGRGGHGARGSPSFRPQNRLRLHLPQPPVQYQYEPSAPS 60
Db 1 MMQGNCTCHMSFHPGRGCPGRGGHGARGSPSFRPQNRLRLHLPQPPVQYQYEPSAPS 60
QY 61 TTFSNSPAENFLPPRPDVFPPPPMPPSAQGLPPCPIRPPPNHOMRHPFVPPCFPPM 120
Db 61 TTFSNSPAENFLPPRPDVFPPPPMPPSAQGLPPCPIRPPPNHOMRHPFVPPCFPPM 120
QY 121 PPMPCPNPPVPVGPAGPGGTPPFMMPPPSMHPHPPVPMQOVNYQYPPGYSHHFPPP 180
Db 121 PPMPCPNPPVPVGPAGPGGTPPFMMPPPSMHPHPPVPMQOVNYQYPPGYSHHFPPP 180
QY 181 SFNSFQNNPSSFLPSANNSSPHFRHLPPYPLPKAPSERRSERLKHDDHRHDSHGR 240
Db 181 SFNSFQNNPSSFLPSANNSSPHFRHLPPYPLPKAPSERRSERLKHDDHRHDSHGR 240
QY 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTPSRHSYERSERERERHHRNRRS 300
Db 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTPSRHSYERSERERERHHRNRRS 300
QY 301 PSLSRYKKEYKSGSYGLSVVPEAGCTPELPGHIIKNTDSWAPPLIIVNHRSPSREK 360
Db 301 PSLSRYKKEYKSGSYGLSVVPEAGCTPELPGHIIKNTDSWAPPLIIVNHRSPSREK 360
QY 361 KQARWEEKDRWSDNQGSGKDKNYTSIKKEPPEETMPDKNEEBEELLKPVMIRCTHSEN 420
Db 361 KQARWEEKDRWSDNQGSGKDKNYTSIKKEPPEETMPDKNEEBEELLKPVMIRCTHSEN 420
QY 421 YYSDDPMDQVGDSTVVGTSLRLDLYDKFEEELGSRQEKAKAAPPPPEPKTKLDELESS 480
Db 421 YYSDDPMDQVGDSTVVGTSLRLDLYDKFEEELGSRQEKAKAAPPPPEPKTKLDELESS 480
QY 481 SSESCESEDSTCSSSDSEVFDVIAEIKRKAHPDRLHDELWYNDPGQWNGDGLCKCSA 540
Db 481 SSESCESEDSTCSSSDSEVFDVIAEIKRKAHPDRLHDELWYNDPGQWNGDGLCKCSA 540
QY 541 KARTGIRHSIYPGEBAIKPCRMTNNAAGRLFHYRITVSPPTNFLTDRPTVIYDDHEIYI 600
Db 541 KARTGIRHSIYPGEBAIKPCRMTNNAAGRLFHYRITVSPPTNFLTDRPTVIYDDHEIYI 600
QY 601 FEGFSMFAHAPLTNIPLCKVIRENIDYTHFIEEMMPENFCVKGLELSFLFRDILELY 660
Db 601 FEGFSMFAHAPLTNIPLCKVIRENIDYTHFIEEMMPENFCVKGLELSFLFRDILELY 660
```



```
QY 661 DWNKGLPFDSPCCPRFHFMPFRFVFLPDGKVELSMHQILLYLLRCSKALVPBEEIA 720
DB 661 DWNKGLPFDSPCCPRFHFMPFRFVFLPDGKVELSMHQILLYLLRCSKALVPBEEIA 720
QY 721 NMLQWEELEWQKVAECKGMI VTNPGTKPSSVRIDQLDREQFNPDTITPPIIVHFGIRPA 780
DB 721 NMLQWEELEWQKVAECKGMI VTNPGTKPSSVRIDQLDREQFNPDTITPPIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKWLKSVYKLRHLANSVKVQTDKQKLAQREBALQKIROKNTMRREVTV 840
DB 781 QLSYAGDPQYQKWLKSVYKLRHLANSVKVQTDKQKLAQREBALQKIROKNTMRREVTV 840
QY 841 ELSQGFMTKGI RSDVCOHAWMLPVLTTHIRYHQCMLHLDKLGITFQDRCLLQLAATHP 900
DB 841 ELSQGFMTKGI RSDVCOHAWMLPVLTTHIRYHQCMLHLDKLGITFQDRCLLQLAATHP 900
QY 901 SHLNFQMPDHARNLSNCGIRQPKYGDVKVHHMMRKKGINTLINIMSLRGQDDTPPS 960
DB 901 SHLNFQMPDHARNLSNCGIRQPKYGDVKVHHMMRKKGINTLINIMSLRGQDDTPPS 960
QY 961 RINHNERLEFLGDVAVVEFLTSTVHLXYLFPSEEGGLATYRTAIVQNOHLAMLAKELEDP 1020
DB 961 RINHNERLEFLGDVAVVEFLTSTVHLXYLFPSEEGGLATYRTAIVQNOHLAMLAKELEDP 1020
QY 1021 FMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEEAQKQFGRLLFNDPDLREVWLN 1080
DB 1021 FMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEEAQKQFGRLLFNDPDLREVWLN 1080
QY 1081 YPLHPLQLQBPNTDROLIETSPVLQKLTPEPEAIGVIFTHVRLARAFRTVGFNHLTL 1140
DB 1081 YPLHPLQLQBPNTDROLIETSPVLQKLTPEPEAIGVIFTHVRLARAFRTVGFNHLTL 1140
QY 1141 GHNORMEFLGDSIMQLVATEYLFHFPDHEHGLTLRLSSLVNVRTQAKVAEELGMQEYA 1200
DB 1141 GHNORMEFLGDSIMQLVATEYLFHFPDHEHGLTLRLSSLVNVRTQAKVAEELGMQEYA 1200
QY 1201 ITNDKTKRPVGLRTKTLADLLESFIAALYTDKOLEYVHTPMNVCFPPRLKEFTLNQDWD 1260
DB 1201 ITNDKTKRPVGLRTKTLADLLESFIAALYTDKOLEYVHTPMNVCFPPRLKEFTLNQDWD 1260
QY 1261 PKSQLOQCCLTLTEGKEPDIPLYKLTQVGPSHARTYTVAVYFKGERIGCGKPSIQQA 1320
DB 1261 PKSQLOQCCLTLTEGKEPDIPLYKLTQVGPSHARTYTVAVYFKGERIGCGKPSIQQA 1320
QY 1321 EMGAAMDALEKYNFPQMAHQKRFGRKYRQELKEMRWERHOREPDETDIKK 1374
DB 1321 EMGAAMDALEKYNFPQMAHQKRFGRKYRQELKEMRWERHOREPDETDIKK 1374
```

RESULT 3

```
US-10-774-974-2
; Sequence 2, Application US/10774974
; Publication No. US20040126867A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/10/774,974
; CURRENT FILING DATE: 2004-02-09
; PRIOR APPLICATION NUMBER: US/09/900,425B
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-774-974-2
```

Query Match 100.0%; Score 7500; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MMOGNTCHRMSPFPGRCPRGCGHGRSAPSFRPQNLELLHPQQPVOYQVEPSPAPS 60
DB 1 MMOGNTCHRMSPFPGRCPRGCGHGRSAPSFRPQNLELLHPQQPVOYQVEPSPAPS 60
QY 61 TTFESNPAPNLFPRDFVFPFPPMPPSAQGLPPCPPIRPPFNNHQRHPPFPPPCFPMP 120
DB 61 TTFESNPAPNLFPRDFVFPFPPMPPSAQGLPPCPPIRPPFNNHQRHPPFPPPCFPMP 120
QY 121 PPMPCPNPPVPFAGPGQGTFFPMMPPPSMMPHPPPPPPVMPQQVNYQYPPGYSHNPPPP 180
DB 121 PPMPCPNPPVPFAGPGQGTFFPMMPPPSMMPHPPPPPPVMPQQVNYQYPPGYSHNPPPP 180
QY 181 SFNSFQNNPSSFLPSANNSSPHFRHLPPYPLKAPSERSSPERLKHYYDHRHRDHSGR 240
DB 181 SFNSFQNNPSSFLPSANNSSPHFRHLPPYPLKAPSERSSPERLKHYYDHRHRDHSGR 240
QY 241 GERHSLDRRERGRSPRRRRQDSRYRSDYDRGRTPSRHRYSRERERERHRHRDNRSS 300
DB 241 GERHSLDRRERGRSPRRRRQDSRYRSDYDRGRTPSRHRYSRERERERHRHRDNRSS 300
QY 301 PSLSRSYKKEYKRSGRSYGLSVVPEPAGCTPELPGEBIIKNVDSWAPPLEIVNHRSPREK 360
DB 301 PSLSRSYKKEYKRSGRSYGLSVVPEPAGCTPELPGEBIIKNVDSWAPPLEIVNHRSPREK 360
QY 361 KRARWEEBKDRWSDNOSGKDKNYTTSIKKEPEETMPDKNEEEEBELIKPWIRCTHSEN 420
DB 361 KRARWEEBKDRWSDNOSGKDKNYTTSIKKEPEETMPDKNEEEEBELIKPWIRCTHSEN 420
QY 421 YTSDDPMDQVGDSTVVTGTSRLRLDYDKFBEELSGROEKAKAAPPPWEPPTKLDDELESS 480
DB 421 YTSDDPMDQVGDSTVVTGTSRLRLDYDKFBEELSGROEKAKAAPPPWEPPTKLDDELESS 480
QY 481 SESECEDESDSTCSSSDSEVFDVIAIKRKAHPDLRDLHDELWYNDPQMGNDGPKCKSA 540
DB 481 SESECEDESDSTCSSSDSEVFDVIAIKRKAHPDLRDLHDELWYNDPQMGNDGPKCKSA 540
QY 541 KARTGIRHSIYPGEBAIKPCRPMNTNAGRLFHYRTVSPPTNFLTDRPTVIYDDHEIYI 600
DB 541 KARTGIRHSIYPGEBAIKPCRPMNTNAGRLFHYRTVSPPTNFLTDRPTVIYDDHEIYI 600
QY 601 FEGFSMFAHAPLTNIPLCKVIRENIDYTHIFIEBMPENFCVKGLELFLFLFDILELY 660
DB 601 FEGFSMFAHAPLTNIPLCKVIRENIDYTHIFIEBMPENFCVKGLELFLFLFDILELY 660
QY 661 DWNKGLPFDSPCCPRFHFMPFRFVFLPDGKVELSMHQILLYLLRCSKALVPBEEIA 720
DB 661 DWNKGLPFDSPCCPRFHFMPFRFVFLPDGKVELSMHQILLYLLRCSKALVPBEEIA 720
QY 721 NMLQWEELEWQKVAECKGMI VTNPGTKPSSVRIDQLDREQFNPDTITPPIIVHFGIRPA 780
DB 721 NMLQWEELEWQKVAECKGMI VTNPGTKPSSVRIDQLDREQFNPDTITPPIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKWLKSVYKLRHLANSVKVQTDKQKLAQREBALQKIROKNTMRREVTV 840
DB 781 QLSYAGDPQYQKWLKSVYKLRHLANSVKVQTDKQKLAQREBALQKIROKNTMRREVTV 840
QY 841 ELSQGFMTKGI RSDVCOHAWMLPVLTTHIRYHQCMLHLDKLGITFQDRCLLQLAATHP 900
DB 841 ELSQGFMTKGI RSDVCOHAWMLPVLTTHIRYHQCMLHLDKLGITFQDRCLLQLAATHP 900
QY 901 SHLNFQMPDHARNLSNCGIRQPKYGDVKVHHMMRKKGINTLINIMSLRGQDDTPPS 960
DB 901 SHLNFQMPDHARNLSNCGIRQPKYGDVKVHHMMRKKGINTLINIMSLRGQDDTPPS 960
QY 961 RINHNERLEFLGDVAVVEFLTSTVHLXYLFPSEEGGLATYRTAIVQNOHLAMLAKELEDP 1020
DB 961 RINHNERLEFLGDVAVVEFLTSTVHLXYLFPSEEGGLATYRTAIVQNOHLAMLAKELEDP 1020
QY 1021 FMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEEAQKQFGRLLFNDPDLREVWLN 1080
DB 1021 FMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEEAQKQFGRLLFNDPDLREVWLN 1080
```

QY 1081 YPLHPLQLOEPNTDRQLIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
Db 1081 YPLHPLQLOEPNTDRQLIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
QY 1141 GHNQRMFLGDSIMQLVATEYLFIHPDPDHHEGLTLRLSSLVNRTQAKVABELGHOEYA 1200
Db 1141 GHNQRMFLGDSIMQLVATEYLFIHPDPDHHEGLTLRLSSLVNRTQAKVABELGHOEYA 1200
QY 1201 ITNDTKRPVGLRTKTLADLLESFIAALYTDKLEYVHTFMNVCFPPRLKEFTLNQDWND 1260
Db 1201 ITNDTKRPVGLRTKTLADLLESFIAALYTDKLEYVHTFMNVCFPPRLKEFTLNQDWND 1260
QY 1261 PKSLOQCCCLTLTEGKEPDIPLYKTLQTVGPHARTYTVAVYFKGERICGKGPSIQQA 1320
Db 1261 PKSLOQCCCLTLTEGKEPDIPLYKTLQTVGPHARTYTVAVYFKGERICGKGPSIQQA 1320
QY 1321 EMGAAMDALEKYNFPQMAHOKRFIGRKYQELKEMWREHOREPDEDEDIKK 1374
Db 1321 EMGAAMDALEKYNFPQMAHOKRFIGRKYQELKEMWREHOREPDEDEDIKK 1374

RESULT 4

US-10-805-919-2
; Sequence 2, Application US/10805919
; Publication No. US20040175828A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/10/805,919
; CURRENT FILING DATE: 2004-03-22
; PRIOR APPLICATION NUMBER: US/09/900,425
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-805-919-2

Query Match 100.0%; Score 7500; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMQGNTRCHMSFPHGRCPRGCGHGHGARSAPSFRPQNLLHLPQQPPVOYQYEPSPAPS 60
Db 1 MMQGNTRCHMSFPHGRCPRGCGHGHGARSAPSFRPQNLLHLPQQPPVOYQYEPSPAPS 60
QY 61 TTFNSPAPNPLPRDDFVFPFPPMPPSAQGLPPCPIRPPFPHQMRHPFPVPPCFPPM 120
Db 61 TTFNSPAPNPLPRDDFVFPFPPMPPSAQGLPPCPIRPPFPHQMRHPFPVPPCFPPM 120
QY 121 PPMPCPNPPVPGAPPGQGTFFPMPMPSPMHPPPPVPVMPQVNYQYPPGYSHNFPPP 180
Db 121 PPMPCPNPPVPGAPPGQGTFFPMPMPSPMHPPPPVPVMPQVNYQYPPGYSHNFPPP 180
QY 181 SFNSFQNNPSSFLPSANNSSPHRLLPYPLPKAPSERSPERLKHDDHRRDHSGR 240
Db 181 SFNSFQNNPSSFLPSANNSSPHRLLPYPLPKAPSERSPERLKHDDHRRDHSGR 240
QY 241 GERHSLDRERGRSPDRRQDSYRSDYDRGTPSRHRSYERSRERERHRDRNR 300
Db 241 GERHSLDRERGRSPDRRQDSYRSDYDRGTPSRHRSYERSRERERHRDRNR 300
QY 301 PSLSRYKKEYKRSGRSYGLSVVPEPAGCTPELPGBIKNKTDSPAPPELVNHRSPSREK 360

Db 301 PSLSRYKKEYKRSGRSYGLSVVPEPAGCTPELPGBIKNKTDSPAPPELVNHRSPSREK 360
QY 361 KEARWEEKDRWSDNOSGKDKNYTISIKEKEPEETMPDKNEEEELLKPVWIRCTHSEN 420
Db 361 KEARWEEKDRWSDNOSGKDKNYTISIKEKEPEETMPDKNEEEELLKPVWIRCTHSEN 420
QY 421 YSSSDPMDQVGSSTVVTGTSRLRDLVDKPFEEELGSRQEKAKAAPPPWEPKTKLDELESS 480
Db 421 YSSSDPMDQVGSSTVVTGTSRLRDLVDKPFEEELGSRQEKAKAAPPPWEPKTKLDELESS 480
QY 481 SSECSDSDSTCSSSDSEVDFVIAEIKRKAHPORLHDELWYNDPQGNNDGFLCKCSA 540
Db 481 SSECSDSDSTCSSSDSEVDFVIAEIKRKAHPORLHDELWYNDPQGNNDGFLCKCSA 540
QY 541 KARTGIRHSIYPGEEAIKPCRPMTNAGRLPHYRITVSPPTNFLTDRPTVIEYDDHEYI 600
Db 541 KARTGIRHSIYPGEEAIKPCRPMTNAGRLPHYRITVSPPTNFLTDRPTVIEYDDHEYI 600
QY 601 FEGFSMFAHAPLTNIPLCKVIRFNIDYTHFTTEEMMPENFCVKGLELFSLFRDILELY 660
Db 601 FEGFSMFAHAPLTNIPLCKVIRFNIDYTHFTTEEMMPENFCVKGLELFSLFRDILELY 660
QY 661 DWNLKGFLFEDSPCCPRFHFMPRFVFLPDGSGEVLSMHQILLYLLRCSKALVPSEEIA 720
Db 661 DWNLKGFLFEDSPCCPRFHFMPRFVFLPDGSGEVLSMHQILLYLLRCSKALVPSEEIA 720
QY 721 NMLQWEELEWQYABECKGMIVTNGTKSSVRIQDLDREQFNPDVITPPIIVHFGIRPA 780
Db 721 NMLQWEELEWQYABECKGMIVTNGTKSSVRIQDLDREQFNPDVITPPIIVHFGIRPA 780
QY 781 QLSYAGDPOYQKLWKSIVKLRHLLANSPKVKTOKDLAQREALQKIQKNTMRREVTV 840
Db 781 QLSYAGDPOYQKLWKSIVKLRHLLANSPKVKTOKDLAQREALQKIQKNTMRREVTV 840
QY 841 ELSQSGFWKTGIRSDVCOHAMLPVLTHIRYHQCIMHLDKLIIGYTFQDRCLLQALWTHP 900
Db 841 ELSQSGFWKTGIRSDVCOHAMLPVLTHIRYHQCIMHLDKLIIGYTFQDRCLLQALWTHP 900
QY 901 SHLNFGMNPDHARNSLNCGRQPKYGRKVVHMMRMKKGINTLINISRLGDDPTPS 960
Db 901 SHLNFGMNPDHARNSLNCGRQPKYGRKVVHMMRMKKGINTLINISRLGDDPTPS 960
QY 961 RINHNERLEFLGDVAVVEFLTSVHLVYLPFSLBEGGLATYRTAIVQNHMLAKKLELDP 1020
Db 961 RINHNERLEFLGDVAVVEFLTSVHLVYLPFSLBEGGLATYRTAIVQNHMLAKKLELDP 1020
QY 1021 FMLYAHGPDLGRESDLRHAMANCFEALIGAVYLEGSLAEAKQLFGRLLFNDPDLREVWLN 1080
Db 1021 FMLYAHGPDLGRESDLRHAMANCFEALIGAVYLEGSLAEAKQLFGRLLFNDPDLREVWLN 1080
QY 1081 YPLHPLQLOEPNTDRQLIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
Db 1081 YPLHPLQLOEPNTDRQLIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
QY 1141 GHNQRMFLGDSIMQLVATEYLFIHPDPDHHEGLTLRLSSLVNRTQAKVABELGHOEYA 1200
Db 1141 GHNQRMFLGDSIMQLVATEYLFIHPDPDHHEGLTLRLSSLVNRTQAKVABELGHOEYA 1200
QY 1201 ITNDTKRPVGLRTKTLADLLESFIAALYTDKLEYVHTFMNVCFPPRLKEFTLNQDWND 1260
Db 1201 ITNDTKRPVGLRTKTLADLLESFIAALYTDKLEYVHTFMNVCFPPRLKEFTLNQDWND 1260
QY 1261 PKSLOQCCCLTLTEGKEPDIPLYKTLQTVGPHARTYTVAVYFKGERICGKGPSIQQA 1320
Db 1261 PKSLOQCCCLTLTEGKEPDIPLYKTLQTVGPHARTYTVAVYFKGERICGKGPSIQQA 1320
QY 1321 EMGAAMDALEKYNFPQMAHOKRFIGRKYQELKEMWREHOREPDEDEDIKK 1374
Db 1321 EMGAAMDALEKYNFPQMAHOKRFIGRKYQELKEMWREHOREPDEDEDIKK 1374

```
US-11-001-993-2
; Sequence 2, Application US/11001993
; Publication No. US20050159384A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS0003-104 (ISIS-5030US.D1)
; CURRENT APPLICATION NUMBER: US/11/001,993
; PRIOR FILING DATE: 2004-12-02
; PRIOR APPLICATION NUMBER: 10/079,185
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-001-993-2

Query Match      100.0%; Score 7500; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMQGNCTHRMSFHPGRCGRGCHGARSAPSFRQNLRLHLPQPPVQYQVEPPSAPS 60
DB 1 MMQGNCTHRMSFHPGRCGRGCHGARSAPSFRQNLRLHLPQPPVQYQVEPPSAPS 60
QY 61 TTFNSPAPNELPPRRDFVPPPPMPPSAGPLPCCPIRPPPNHQMRRPVPVPCFPPM 120
DB 61 TTFNSPAPNELPPRRDFVPPPPMPPSAGPLPCCPIRPPPNHQMRRPVPVPCFPPM 120
QY 121 PPPMPCNNPPVPGAPGGQGTFFMMPPPSMHPPPPVPVMPQVNYQYPPGYSHNFPPP 180
DB 121 PPPMPCNNPPVPGAPGGQGTFFMMPPPSMHPPPPVPVMPQVNYQYPPGYSHNFPPP 180
QY 181 SFNSFQNNPSSFLPSANSSSHFRLHPYPLPKAPSERRSPRLKHYDDHRRDHSGR 240
DB 181 SFNSFQNNPSSFLPSANSSSHFRLHPYPLPKAPSERRSPRLKHYDDHRRDHSGR 240
QY 241 GERHSLDRRGRSPDRRQDSRYSDYDRGTPSRHSYSRERERERHRDRNRSS 300
DB 241 GERHSLDRRGRSPDRRQDSRYSDYDRGTPSRHSYSRERERERHRDRNRSS 300
QY 301 PSLSRSYKKEYKRSRSGYGLSVVPEPAGCTPELPGEIINKTDSWAPPLPPIVNHRSRREK 360
DB 301 PSLSRSYKKEYKRSRSGYGLSVVPEPAGCTPELPGEIINKTDSWAPPLPPIVNHRSRREK 360
QY 361 KQARWEEKDRWSDNOSGKDKNYTSIKKKEPEETMPDKNEEBEELLPKWIRCTHSEN 420
DB 361 KQARWEEKDRWSDNOSGKDKNYTSIKKKEPEETMPDKNEEBEELLPKWIRCTHSEN 420
QY 421 YVSSDPMQVGSSTVVGTSRLRLDYDKPFEELGSRQEKAKAAPPPWPKTKLDELESS 480
DB 421 YVSSDPMQVGSSTVVGTSRLRLDYDKPFEELGSRQEKAKAAPPPWPKTKLDELESS 480
QY 481 SESECESDSDTSCSSSDSEVFVDIAEIKRKAHPDRLHDELWYNDPGQMDGCLCKCSA 540
DB 481 SESECESDSDTSCSSSDSEVFVDIAEIKRKAHPDRLHDELWYNDPGQMDGCLCKCSA 540
QY 541 KARTGIRHSIYPGEBAIKPCRPMTNAGRLFHYRITVSPPTNFLTDRPTVIEYDDHEVI 600
DB 541 KARTGIRHSIYPGEBAIKPCRPMTNAGRLFHYRITVSPPTNFLTDRPTVIEYDDHEVI 600
QY 601 PFGFSMFAHAPLTNIPCLKVIREFNIDYTHFIEEMMPENFCVKGLELFLSLFLFRDILELY 660
DB 601 PFGFSMFAHAPLTNIPCLKVIREFNIDYTHFIEEMMPENFCVKGLELFLSLFLFRDILELY 660
QY 661 DWNLKGPLFEDSPCCPRFHFMPRFRVFLPDGQKEVLSMHQIILLYLLRCSKALVPBEEIA 720
DB 661 DWNLKGPLFEDSPCCPRFHFMPRFRVFLPDGQKEVLSMHQIILLYLLRCSKALVPBEEIA 720
QY 721 NMLQWEELEWQKYABECKGMIVTNPGTKPSSSVRIDQDRQFNPDPVITFPPIIVHFGIRPA 780
DB 721 NMLQWEELEWQKYABECKGMIVTNPGTKPSSSVRIDQDRQFNPDPVITFPPIIVHFGIRPA 780
QY 781 QLSYAGDPQOYQKLWKSYYVKLRHLLANS PKVQKQTDKQKLAQREBALQKIROKNTMRREVTV 840
DB 781 QLSYAGDPQOYQKLWKSYYVKLRHLLANS PKVQKQTDKQKLAQREBALQKIROKNTMRREVTV 840
QY 841 ELSQGFQWTKGIRSDVCQHAMMLPVLTHIRYHQCIMHLDKLIIGYTFQDRCLLOLANTHP 900
DB 841 ELSQGFQWTKGIRSDVCQHAMMLPVLTHIRYHQCIMHLDKLIIGYTFQDRCLLOLANTHP 900
QY 901 SHHLNFGMPPDHARNSLSNCGIRQPKYGDVKVHHMHRKKGINTLINIMSRIGQDDPTPS 960
DB 901 SHHLNFGMPPDHARNSLSNCGIRQPKYGDVKVHHMHRKKGINTLINIMSRIGQDDPTPS 960
QY 961 RINHNERLEFLGDVAVVEFLTSHVLYLFPFSLBEGGLATYRTAIVQNHLMAMLAKCLELDP 1020
DB 961 RINHNERLEFLGDVAVVEFLTSHVLYLFPFSLBEGGLATYRTAIVQNHLMAMLAKCLELDP 1020
QY 1021 FMLYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLEBAKQLFGRLLFNDPDLREVWLN 1080
DB 1021 FMLYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLEBAKQLFGRLLFNDPDLREVWLN 1080
QY 1081 YPLHPLQLOQEPNTDROLIETSPVLQKLTFFEEAIGVIFTHVRLARAFRTLVTFGNHLLT 1140
DB 1081 YPLHPLQLOQEPNTDROLIETSPVLQKLTFFEEAIGVIFTHVRLARAFRTLVTFGNHLLT 1140
QY 1141 GHNRQMEFLGDSIMQLVATEYLFIHPDHHHEGLTLRLSSLVNNRTQAKVAEELGMOEYA 1200
DB 1141 GHNRQMEFLGDSIMQLVATEYLFIHPDHHHEGLTLRLSSLVNNRTQAKVAEELGMOEYA 1200
QY 1201 IYNDTKRPVGLRTKTLADLLSFAALYTDKDLVYHVTFMNVCFPPRLKEPILNQDND 1260
DB 1201 IYNDTKRPVGLRTKTLADLLSFAALYTDKDLVYHVTFMNVCFPPRLKEPILNQDND 1260
QY 1261 PKSQLOQCCLTARTGKEPDIPLYKTLQTVGSHARTYTVAVYFKGERGCGKGPSIQQA 1320
DB 1261 PKSQLOQCCLTARTGKEPDIPLYKTLQTVGSHARTYTVAVYFKGERGCGKGPSIQQA 1320
QY 1321 EMGAAMDALKENYFPOMAHQKRFGRKYRQELKEMWREHREHREDEDEDIKK 1374
DB 1321 EMGAAMDALKENYFPOMAHQKRFGRKYRQELKEMWREHREHREDEDEDIKK 1374

RESULT 6
US-10-723-860-281
; Sequence 281, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 281
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-281
```

Query Match		99.8%;	Score 7486;	DB 5;	Length 1374;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 1372;		Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
Qy	1	MMQGNCTCHRMSPHGRGCGRGRGGHGCARPSAFRPNQLRLHLPQPPVQYQYEPSPAPS	60			
Db	1	MMQGNCTCHRMSPHGRGCGRGRGGHGCARPSAFRPNQLRLHLPQPPVQYQYEPSPAPS	60			
Qy	61	TTFSNSPAPNFIPLPRDFVFPFPPMPPSAQGLPPCPIRPPPPNHQMRHPFPVPPCFPPM	120			
Db	61	TTFSNSPAPNFIPLPRDFVFPFPPMPPSAQGLPPCPIRPPPPNHQMRHPFPVPPCFPPM	120			
Qy	121	PPMPCPNPPVPGAPGGGTFFMMPPPSMHPPPPVMQOVNYYPGYSHHNFPFP	180			
Db	121	PPMPCPNPPVPGAPGGGTFFMMPPPSMHPPPPVMQOVNYYPGYSHHNFPFP	180			
Qy	181	SFNSFQNNPSSFLPSANNSSPHFRLLPPYPLPKAPSERSSPERLKHYYDDHRRDHSRGR	240			
Db	181	SFNSFQNNPSSFLPSANNSSPHFRLLPPYPLPKAPSERSSPERLKHYYDDHRRDHSRGR	240			
Qy	241	GERHSLDRRGRSPDRRRQDSRYSDYDRGRTPSRHSYERSRERERHRHRDNRSS	300			
Db	241	GERHSLDRRGRSPDRRRQDSRYSDYDRGRTPSRHSYERSRERERHRHRDNRSS	300			
Qy	301	PSLSYKKEYKRSGRSYGLSVVPEPAGCTPELPGELIKNTDSWAPPLEIVNHRSPSREK	360			
Db	301	PSLSYKKEYKRSGRSYGLSVVPEPAGCTPELPGELIKNTDSWAPPLEIVNHRSPSREK	360			
Qy	361	KXARWEEEDRWSDNOSGKDKNYTSGIKEPEETMPDKNEEEELLKPVWIRCTHSEN	420			
Db	361	KXARWEEEDRWSDNOSGKDKNYTSGIKEPEETMPDKNEEEELLKPVWIRCTHSEN	420			
Qy	421	YSSSDPMDQVGDSTVVGTSLRLDLYDKFBEELSGRQEKAKAAPPPWEPKTKLDEPLESS	480			
Db	421	YSSSDPMDQVGDSTVVGTSLRLDLYDKFBEELSGRQEKAKAAPPPWEPKTKLDEPLESS	480			
Qy	481	SESECEDESDSTCSSSDSEVFVIAIEIKKKAHPDRLHDELWYNDPGQWMDGLCKCSA	540			
Db	481	SESECEDESDSTCSSSDSEVFVIAIEIKKKAHPDRLHDELWYNDPGQWMDGLCKCSA	540			
Qy	541	KARTGIRUSIYIGBEAIPKCRPMTNNAAGLPHYRITVSPPTNFTLDRPTVIEYDDHEVI	600			
Db	541	KARTGIRUSIYIGBEAIPKCRPMTNNAAGLPHYRITVSPPTNFTLDRPTVIEYDDHEVI	600			
Qy	601	FEGFSMFAHAPLNIPLCKVIRENIDYTIHFIEEMMPENFCVKGLELFSLFRDILELY	660			
Db	601	FEGFSMFAHAPLNIPLCKVIRENIDYTIHFIEEMMPENFCVKGLELFSLFRDILELY	660			
Qy	661	DNMLKGPLFEDSPCCPRHFMPFRVFLPDGGSKEVLSMHQIILLYLLRCSKALVPEETA	720			
Db	661	DNMLKGPLFEDSPCCPRHFMPFRVFLPDGGSKEVLSMHQIILLYLLRCSKALVPEETA	720			
Qy	721	NMLQWELEWQYAECKGMIVNPGTKPSSSVRIDOLDRQFNPDVITTPIIVHFGIRPA	780			
Db	721	NMLQWELEWQYAECKGMIVNPGTKPSSSVRIDOLDRQFNPDVITTPIIVHFGIRPA	780			
Qy	781	QLSYAGDPQYQKLWKSYYKLRLHLLANS PKVKOTDKQKLAQREBALQKIQKNTMRRETV	840			
Db	781	QLSYAGDPQYQKLWKSYYKLRLHLLANS PKVKOTDKQKLAQREBALQKIQKNTMRRETV	840			
Qy	841	ELSSQGFWKTGIRSDVCQHAWMLPVLTTHIRYHQCILMHLDKLIGYTFQDRCLLQLANTHP	900			
Db	841	ELSSQGFWKTGIRSDVCQHAWMLPVLTTHIRYHQCILMHLDKLIGYTFQDRCLLQLANTHP	900			
Qy	901	SHHLNFGMNDHARNISLNCIGIQPKYGDGRKVHMMRKKGINTLNINMSRLGQDDPTPS	960			
Db	901	SHHLNFGMNDHARNISLNCIGIQPKYGDGRKVHMMRKKGINTLNINMSRLGQDDPTPS	960			
Qy	961	RINHNERLEFLGDAVVEFLTSVHLYYLFPSLEBGGATYRTAIVQNHQHLAMLAKKLELDP	1020			
Db	961	RINHNERLEFLGDAVVEFLTSVHLYYLFPSLEBGGATYRTAIVQNHQHLAMLAKKLELDP	1020			

RESULT 7

US-10-756-149-4809
; Sequence 4809, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:

; APPLICANT: Aziz, Natasha

; APPLICANT: Zlotnik, Albert

; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND

; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER

; FILE REFERENCE: file

; CURRENT APPLICATION NUMBER: US/10/756,149

; CURRENT FILING DATE: 2004-01-12

; NUMBER OF SEQ ID NOS: 5818

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 4809

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-756-149-4809

Query Match 99.8%; Score 7486; DB 5; Length 1374;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MMQGNCTCHRMSPHGRGCGRGRGGHGCARPSAFRPNQLRLHLPQPPVQYQYEPSPAPS 60

Db 1 MMQGNCTCHRMSPHGRGCGRGRGGHGCARPSAFRPNQLRLHLPQPPVQYQYEPSPAPS 60

Qy 61 TTFSNSPAPNFIPLPRDFVFPFPPMPPSAQGLPPCPIRPPPPNHQMRHPFPVPPCFPPM 120

Db 61 TTFSNSPAPNFIPLPRDFVFPFPPMPPSAQGLPPCPIRPPPPNHQMRHPFPVPPCFPPM 120

Qy 121 PPMPCPNPPVPGAPGGGTFFMMPPPSMHPPPPVMQOVNYYPGYSHHNFPFP 180

Db 121 PPMPCPNPPVPGAPGGGTFFMMPPPSMHPPPPVMQOVNYYPGYSHHNFPFP 180

Qy 181 SFNSFQNNPSSFLPSANNSSPHFRLLPPYPLPKAPSERSSPERLKHYYDDHRRDHSRGR 240

Db 181 SFNSFQNNPSSFLPSANNSSPHFRLLPPYPLPKAPSERSSPERLKHYYDDHRRDHSRGR 240

Qy 241 GERHSLDRRGRSPDRRRQDSRYSDYDRGRTPSRHSYERSRERERHRHRDNRSS 300

Db 241 GERHSLDRRGRSPDRRRQDSRYSDYDRGRTPSRHSYERSRERERHRHRDNRSS 300

Qy 301 PSLSYKKEYKRSGRSYGLSVVPEPAGCTPELPGELIKNTDSWAPPLEIVNHRSPSREK 360

Db 301 PSLSYKKEYKRSGRSYGLSVVPEPAGCTPELPGELIKNTDSWAPPLEIVNHRSPSREK 360

361 KRARWEEKDRWSDNOSGDKXNTS1KEKEPETMPDKNEEBEELKKPVWIRCTHSEN 420
361 KRARWEEKDRWSDNOSGDKXNTS1KEKEPETMPDKNEEBEELKKPVWIRCTHSEN 420
421 YSSDPMDQVGDSTVVTGTSRLDLYDKFEELSGRQEKAKAAPPPWPKTKLDEDESS 480
421 YSSDPMDQVGDSTVVTGTSRLDLYDKFEELSGRQEKAKAAPPPWPKTKLDEDESS 480
481 SESECEDESDSTCSSSDSEVFVIAEIKRKAHPDRHLDELWYNDPGQNDGPKCKSA 540
481 SESECEDESDSTCSSSDSEVFVIAEIKRKAHPDRHLDELWYNDPGQNDGPKCKSA 540
541 KARTGIRUSIYGEBAIKPCRMTNAGRLFYRITVSPPTNFLTDRPTVIVYDDHEVI 600
541 KARTGIRUSIYGEBAIKPCRMTNAGRLFYRITVSPPTNFLTDRPTVIVYDDHEVI 600
601 FEGFSMAFAHPLTNIPLCKVIRENIDYTHIFIEEMMPENFCVKGLFSLFLFRDILELY 660
601 FEGFSMAFAHPLTNIPLCKVIRENIDYTHIFIEEMMPENFCVKGLFSLFLFRDILELY 660
661 DWNKGLPDPSPCCPRHFMRFRVFLPDGKKEVLSMIOIILLYLLRCSKALVPEEETA 720
661 DWNKGLPDPSPCCPRHFMRFRVFLPDGKKEVLSMIOIILLYLLRCSKALVPEEETA 720
721 NMLQWEELEWQKVAEBCGMIVNPGTKSSVRIDQLDRQFNPDVITPPIIVHFGIRPA 780
721 NMLQWEELEWQKVAEBCGMIVNPGTKSSVRIDQLDRQFNPDVITPPIIVHFGIRPA 780
781 QLSYAGDPQYQKLWKSYYVLRHLLANSVKVQTDKQKLAQREBALQKIRQNTMRREVTV 840
781 QLSYAGDPQYQKLWKSYYVLRHLLANSVKVQTDKQKLAQREBALQKIRQNTMRREVTV 840
841 ELSSQGFMTKIRSDVCQHAMMLPVLTHIRYHQCMLHLDKLIIGYTFQDRCLLQALWTHP 900
841 ELSSQGFMTKIRSDVCQHAMMLPVLTHIRYHQCMLHLDKLIIGYTFQDRCLLQALWTHP 900
901 SHHNLFGMNPDHARNSLSNGIRQPKYQDKRVHMHMRKKGINTLINIMSLRGQDDPTPS 960
901 SHHNLFGMNPDHARNSLSNGIRQPKYQDKRVHMHMRKKGINTLINIMSLRGQDDPTPS 960
961 RINHNERLEFLGDVAVFEFTSVHLYLFPSSLEGGGLATYRTAIVQOHLAMLAKELEDP 1020
961 RINHNERLEFLGDVAVFEFTSVHLYLFPSSLEGGGLATYRTAIVQOHLAMLAKELEDP 1020
1021 FMYLHARGPDLCRSDLRHAMANCFEALIGAVYLEGSLEBAKQLFGRLLFNDPDLREVWLN 1080
1021 FMYLHARGPDLCRSDLRHAMANCFEALIGAVYLEGSLEBAKQLFGRLLFNDPDLREVWLN 1080
1081 YPLHPLQLOQEPNTDROLIETSPVLQKLTREPEEAGVIFTHVRLIARAFTLRTVGFNHLTL 1140
1081 YPLHPLQLOQEPNTDROLIETSPVLQKLTREPEEAGVIFTHVRLIARAFTLRTVGFNHLTL 1140
1141 GHNMREFLGDSIMQVAVTEYLFIHPDDHHEGLTLRLSSLVNNRTOAKVABELGMOEYA 1200
1141 GHNMREFLGDSIMQVAVTEYLFIHPDDHHEGLTLRLSSLVNNRTOAKVABELGMOEYA 1200
1201 ITNDTKRPVGLRTKTLADLLESFIAALYTDKOLEYVHTFMNVCFFPRLKEFLINQDWDND 1260
1201 ITNDTKRPVGLRTKTLADLLESFIAALYTDKOLEYVHTFMNVCFFPRLKEFLINQDWDND 1260
1261 PKSQLQOCCLTLTREGKEPDIPLYKTLQVGPSHARTYTVAVYFKGERICGCGPSIOQA 1320
1261 PKSQLQOCCLTLTREGKEPDIPLYKTLQVGPSHARTYTVAVYFKGERICGCGPSIOQA 1320
1321 EMGAAMDALKEYNFPOWAKHOKRIGRYQELKEMEREWEREDEDEDIKK 1374
1321 EMGAAMDALKEYNFPOWAKHOKRIGRYQELKEMEREWEREDEDEDIKK 1374

RESULT 8

US-11-097-143-2409

; Sequence 2409, Application US/11097143

; Publication No. US2005020858A1

GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2409
; LENGTH: 1327
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-2409

Query Match 39.5%; Score 2963; DB 6; Length 1327;

Best Local Similarity 45.6%; Pred. No. 7.8e-173;

Matches 621; Conservative 213; Mismatches 327; Indels 200; Gaps 32;

QY 90 QGPLPCPTRPPPNHQMHPFPVPCFPMPMPMPMPCPN-NPPVPGAP-----P 137
DB 3 QPLPLPPVQPA-----PPPPPPPEEDLSPPGVGVFHSNYSNESHQ 46
QY 138 GQGTFFPMPPSPMPPPPVMPVQVNYQYPPGY-----SHNFP 179
DB 47 SSKSLDYVPEPAPYASS---VPSYDYPQPPAYGEGYAYNEAQKYGQSSHVQYQY 103
QY 180 PSFNSQNNPSPGFLPSANNSSPHFRLPPYPL-----PKAPSRSPERLKHVDDHHRD 235
DB 104 PA-----SGSFLYES-----YKVPDRYPAYSSNYRPPSERQ-----RYTS 139
QY 236 HSHGCRGHRSLDRRGRSPDRRRQDSYRSYDGRTPSRHSYRSYRSRERERHRHR 295
DB 140 NSSSQGYHYP-----GYSGR-----RYEQRHQ-----EHRQIDSYAHEPRGHY 183
QY 296 DNRSPSLERSYKCKYKRSRSGSYGLSVVPAGCTPELPGEIIKNTDSMAPPLEIVNHS 355
DB 184 AHRQAGSQRGY---YGSARN-----QVSDDYSPRGHERERN 219
QY 356 PSREKKRA--RWEERKDR---WSDNQSGKDKNNTS1KEKPEETMPDKNEEBEELK 409
DB 220 ETLEKTRAKPKVETERDLRLRQWCSNFC-----EKEDYVKKGNALSEADAPV 267
QY 410 PWNIRTHSENYSYSSDPMQVGDSTVVTGTSRLDLYDKFEELSGRQEKAKAAPPPWEP 469
DB 268 ESWVRSPAEIYYERTK---GENEVGRARLQKLTFLDEELQRAKRVREKLPVVP 323
QY 470 KTKL-----DEDLSSSECEDESDSTCSSSDSEVFDV-----IABIKRKAH 514
DB 324 PRKARRRVCKHKHKEACSSSSDDSDDEDA-----FKIBODCCMEELSRKQV 373
QY 515 PDRLHDELWYNDPGQNDGPKCKSAKARRTGIRHSIYPGEBAIKPCRMTNAGRLFYH 574
DB 374 PQRVHADLWINDAGENWDDGFLCCKSAKRRIGIRHGIYFGETGYKCLDPSNNAKGLFYH 433

Qy	575	RITVSPPTNFLTDRPTVIEYDDHXYIPEGSMFAHAPLWNIPLCKVIRFNIDTTHIFEE	634
Db	434	RISISPTNFLTTPPTIKDHEHFEFEGFSLSHVLSDPLCKVIRFNIETIYEEES	493
Qy	635	MPNPNFCVKGLELPSLFLFDILELYDWNLLKGPL-----PDSPPCCPRFHFMRFRVFLP	690
Db	494	KMPENFTIHEDIFPKYLFHLELSDVFNLPNLPNPSGNVEES---CPAFHFFRFRVRLDLP	550
Qy	691	DGGKEVLSMHQILLYLLRCSKALVPBEEIANMLQWEELEWQYABECKGMIVTNPQTKPS	750
Db	551	DNGKEVLAMVEVLYRLLDNSAQLVERQQLLHNLQISOSEWQNYVDFIKGMLVTKPGYKPC	610
Qy	751	SVRIDOLDREQFNDV-----ITFPITIVHEGIRPAQLSVAGDPOYOKLWKSYYKLR	801
Db	611	SLRVDQJDRN--NSDLPPECVDVRETGISHPALVHFGICHPOLSVAGNPEYOKANREYKYR	668
Qy	802	HLANSPKVQTDQKLAQREALQKIRQKNTWREVTVELSSQGFWKTGIRSDVCOHAM	861
Db	669	HLMANMSKPFKDKRKLKEEQRLQEMRTQGRMKRNTVAISSEGFYRTGIMCDVQVHAM	728
Qy	862	MLPVLTHIRYHQCIMHLDKLIQVTFODRCLQLAMTHPSHHLNFGMNPDHARNSLSNCG	921
Db	729	LIPVLTLGHLRPHKSLDLLEESIGYRFKRYLLQLALTHPSYKENGYTNPDPHARNSLTNCG	788
Qy	922	IROPKYGD RKVHHMHMKKGINTLINMSRLGQDDPTPSRINNNERLEFGLDVAWREFLTS	981
Db	789	IROPEYGD RKLIHYNNTRKRGINTLUVSIMS RPKGHEHTVSNITNNERLEFGLDVAWREFLSS	848
Qy	982	VHLYLFPFLEEGGLATYRTAIVQNOHLAMAKKLELDPFMLYAHGPDLCRESDLRHAMA	1041
Db	849	IHLFMPFELEEGGLATYRAAIVQNOHLALLAKKLQLEEFMLYAHGSDLCHELELRHAMA	908
Qy	1042	NCFEALLGAVYLEGSLBEAKQLFGRLLP-NDPDLREVWLYNPLHPLQLOQFNTDRQLIET	1100
Db	909	NCFEALMGALLDGGIIVKVADEVFTDALFRQDEKLLSIWKNLPEHPLOEQEPQLGDRSCIDS	968
Qy	1101	SPVLQKLTPEEEAGVIFTHVRLIKARAFTLRTVGFNHLTLGHQNMREFLGDSIMQLVATE	1160
Db	969	YRVUKELTKEDSICIGFKHIRLLARAFTDRSIGTFHLTIGSNQRLBEFLGDTVLQICSE	1028
Qy	1161	YLFTHFPDHPHGHITLRLSSLVNNRNTQAKVABELGMQSYAI-TNDKTRPVGLRTKTLAD	1219
Db	1029	YLYRHFPEHEGHLSLLRSSLVNNRNTQAVVCDLLGMPKYAVYANPK-----ADLKTDRAD	1084
Qy	1220	LLESFIAALYTDKLEYVHTFMNVCFPRUKERTLNQDNDPKSLOQOCCCLTLRT-EGKE	1278
Db	1085	LLEAFPLGALYVDKGLLYCEOFCHVCLFRQLQFIFMNDQWNDPKSKLOCCCLTLRTMDGGE	1144
Qy	1279	PDIPLYKTLQTVGPSHARTTVAVYFKEGIRIGCGKPSIOQAENGAAMDALKEYN--FPQ	1336
Db	1145	PDIPIYKVVESAGTNTRVYKVAVYFRSKRLATSSGSSIQAENNAAKALENSRDLFPQ	1204
Qy	1337	MAHQQRIGIRKYQEL---KEMRWERBHQREPDETEDIKK	1374
Db	1205	LDHQKRVIAKSIKKQGTGNELNDSDRQHQE-----EKIKR	1239

RESULT 9

```

US-10-205-331-116
; Sequence 116, Application US/10205331
; Publication No. US20040059326A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
;
; TITLE OF INVENTION: Identification a
;
; FILE REFERENCE: WL-A-018199
;
; CURRENT APPLICATION NUMBER: US/10/20
; CURRENT FILING DATE: 2002-07-24
;
; PRIOR APPLICATION NUMBER: GB 0118354
;
; PRIOR FILING DATE: 2001-07-27

```

```

; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 116
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Putative ribonuclease III
US-10-205-331-116

```

Query Match 38.0%; Score 2847; DB 4; Length 541;
Best Local Similarity 99.3%; Pred. No. 3.2e-166;
Matches 537; Conservative 0; Mismatches 4; Indels

Qy	834	MRREVTVELSSQGFWKGTIBSDVCOHAMMLPVLTTHIRYHOCILMHLDKLIGYTFQDRCLL	893
Db	1	MRREVTVELSSQGFWKGTIBSDVCOHAMMLPVLTTHIRYHOCILMHLDKLIGYTFQDRCLL	60
Qy	894	QLAMWTHSHLNFQMPDPHARNSLNSCGIRQPKYGDKRVHMHMRKKGINTLINIMSRLG	953
Db	61	QLAMWTHSHLNFQMPDPHARNSLNSCGIRQPKYGDKRVHMHMRKKGINTLINIMSRLG	120
Qy	954	QDDPTPSRINNHNERLEFLGDVAVVEFTSVHLYYLFPSLEEGGLATYRTAIVQNHMLA	1013
Db	121	QDDPTPSRINNHNERLEFLGDVAVVEFTSVHLYYLFPSLEEGGLATYRTAIVQNHMLA	180
Qy	1014	KXLELDPFMYIYANGPDI.CRESDI.RHAMANCEFALIGAVYLEGSL.EEAKOLFGRLLFNDDP	1073
Db	181	KXLELDRFMYIYANGPDI.CRESDLRHAMANCEFALIGAVYLEGSL.EEAKOLFGRLLFNDDP	240
Qy	1074	LREWMLNYPHPILOQBPNTDRQLIETSPVLQKLTBFEEAIGVIFTHVRLARAFILRTV	1133
Db	241	LREWMLNYPHPILOQBPNTDRQLIETSPVLQKLTBFEEAIGVIFTHVRLARAFILRTV	300
Qy	1134	GFNHLTLGHNORMEFLGDSIMQLVATEYLFTHPPDHHEGHLTLRSSLVNNRTOAKVAEE	1193
Db	301	GFNHLTLGHNORMEFLGDSIMQLVATEYLFTHPPDHHEGHLTLRSSLVNNRTOAKVAEE	360
Qy	1194	LGMQEVAITNDKTRPVGLATKTLADLLESFIAALYTDKDLEYVHPFMNVCFFPRLKEFI	1253
Db	361	LGMQEVAITNDKTRPVALRTKTLADLLESFIAALYTDKDLEYVHTPMNVCFFPRLKEFI	420
Qy	1254	LNQDNDPKSQLOCCCLTLTEGKBPDIPLYKTLQTVGPHSHARTYTVAVYFKGERIGCGK	1313
Db	421	LNQDNDPKSQLOCCCLTLTEGKBPDIPLYKTLQTVGPHSHARTYTVAVYFKGERIGCGK	480
Qy	1314	GPSIQQAEMGAAMDALKYNFPQMAHQKRIIGRKYRQELKEMRWERHQBREPDETEDIK	1373
Db	481	GPSIQQAEMGAAMDALKYNFPQMAHQKRIERKYRQELKEMRWERHQBREPDETEDIK	540
Qy	1374	K 1374	
Db	541	K 541	

RESULT 10

```

RESUB: 10
US-10-774-974-37
; Sequence 37, Application US/10774974
; Publication No. US20040126867A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongliang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And C
; FILE REFERENCE: IS/5029/LS/PH-0522
; CURRENT APPLICATION NUMBER: US/10/774,974
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/900,425B
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 466
; TYPE: PRT

```



```
; ORGANISM: Homo sapiens
US-10-774-974-37

Query Match      32.8%; Score 2463; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 9.9e-143;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 909 NPDHARNLSNCGIRPKYGDVHMHMKKGINLIMSLRGQDDPTPSININERL 968
Db 1 NPDHARNLSNCGIRPKYGDVHMHMKKGINLIMSLRGQDDPTPSININERL 60

QY 969 EFGDVAVEFLTSVHLIYLPSPLEEGGLATYRTAIVQNHLMAMAKKLEDPFMLVAHGP 1028
Db 61 EFGDVAVEFLTSVHLIYLPSPLEEGGLATYRTAIVQNHLMAMAKKLEDPFMLVAHGP 120

QY 1029 DLCRESDLRHAMANCEALIGAVYLEGSLEEAQKLFGRLLFNDPDLREVWLNYPHLPQL 1088
Db 121 DLCRESDLRHAMANCEALIGAVYLEGSLEEAQKLFGRLLFNDPDLREVWLNYPHLPQL 180

QY 1089 QEPNTRQLIETSPVLQKLTPEEAIGVIFTHVRLARAFPLRTVGFNHLTLGHNMREF 1148
Db 181 QEPNTRQLIETSPVLQKLTPEEAIGVIFTHVRLARAFPLRTVGFNHLTLGHNMREF 240

QY 1149 LGSIMQVATEYLFTHFPDHHGHGLTLLESSLVNVRTQAKVAELGMQSYAITNDKTKR 1208
Db 241 LGSIMQVATEYLFTHFPDHHGHGLTLLESSLVNVRTQAKVAELGMQSYAITNDKTKR 300

QY 1209 PVGLRTKTLADLESFIAALYTDKLEYVHTFMNVCFPPRLKEFILNQDNDPKSQLOQC 1268
Db 301 PVGLRTKTLADLESFIAALYTDKLEYVHTFMNVCFPPRLKEFILNQDNDPKSQLOQC 360

QY 1269 CLTLRTGKPEDPIPLYKTLQTVGSPSHARTYTVAVYFKGERIGCGKGPSIQQAEMGA 1328
Db 361 CLTLRTGKPEDPIPLYKTLQTVGSPSHARTYTVAVYFKGERIGCGKGPSIQQAEMGA 420

QY 1329 LEKYNPQMAHQRFGRKRYQBLKEMREREHQPDEDEDIKK 1374
Db 421 LEKYNPQMAHQRFGRKRYQBLKEMREREHQPDEDEDIKK 466

RESULT 11
US-10-103-313-471
; Sequence 471, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ07C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 471
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-471

Query Match      26.1%; Score 1961; DB 4; Length 378;
Best Local Similarity 98.9%; Pred. No. 4.9e-112;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 997 ATYRTAIVQNHLMAMAKKLEDPFMLYAHGPDLCRESDLRHAMANCEALIGAVYLEGS 1056
Db 1 ATYRTAIVQNHLMAMAKKLEDPFMLYAHGPDLCRESDLRHAMANCEALIGAVYLEGS 60

QY 1057 LBEAKQLFGRLLFNDPDLREVWLNYPHLPQLQEPNTRQLIETSPVLQKLTPEEAIGV 1116
Db 61 LBEAKQLFGRLLFNDPDLREVWLNYPHLPQLQEPNTRQLIETSPVLQKLTPEEAIGV 120

QY 1117 IFTHVRLARAFPLRTVGFNHLTLGHNMREFLGSIMQVATEYLFTHFPDHHGHGLTL 1176
Db 1177 IFTHVRLARAFPLRTVGFNHLTLGHNMREFLGSIMQVATEYLFTHFPDHHGHGLTL 180

QY 1177 LRSSLVNNRTQAKVAELGMQSYAITNDKTKRPVGLRTKTLADLESFIAALYTDKLEY 1236
Db 181 LRSSLVNNRTQAKVAELGMQSYAITNDKTKRPVGLRTKTLADLESFIAALYTDKLEY 240

QY 1237 VHTFMNVCFPPRLKEFILNQDNDPKSQLOQCCLTLRTGKPEDPIPLYKTLQTVGSPSHAR 1296
Db 241 VHTFMNVCFPPRLKEFILNQDNDPKSQLOQCCLTLRTGKPEDPIPLYKTLQTVGSPSHAR 300

QY 1297 YTVAVYFKGERIGCGKGPSIQQAEMGAAMDALAEKYNFPQMAHQRFGRKRYQBLKEMR 1356
Db 301 YTVAVYFKGERIGCGKGPSIQQAEMGAAMDALAEKYNFPQMAHQRFGRKRYQBLKEMR 360

QY 1357 WEREHQPDEDEDIKK 1374
Db 361 WEREHQPDEDEDIKK 378

RESULT 12
US-10-103-313-307
; Sequence 307, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ07C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 307
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-307

Query Match      17.6%; Score 1320; DB 4; Length 263;
Best Local Similarity 97.7%; Pred. No. 6.7e-73;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1118 FTHVRLARAFPLRTVGFNHLTLGHNMREFLGSIMQVATEYLFTHFPDHHGHGLTL 1177
Db 7 FTHVRLARAFPLRTVGFNHLTLGHNMREFLGSIMQVATEYLFTHFPDHHGHGLTL 66

QY 1178 RSSLVNNRTQAKVAELGMQSYAITNDKTKRPVGLRTKTLADLESFIAALYTDKLEY 1237
Db 67 RSSLVNNRTQAKVAELGMQSYAITNDKTKRPVGLRTKTLADLESFIAALYTDKLEY 126

QY 1238 HTFMNVCFPPRLKEFILNQDNDPKSQLOQCCLTLRTGKPEDPIPLYKTLQTVGSPSHART 1297
Db 127 HTFMNVCFPPRLKEFILNQDNDPKSQLOQCCLTLRTGKPEDPIPLYKTLQTVGSPSHART 186

QY 1298 YTVAVYFKGERIGCGKGPSIQQAEMGAAMDALAEKYNFPQMAHQRFGRKRYQBLKEMR 1357
Db 187 YTVAVYFKGERIGCGKGPSIQQAEMGAAMDALAEKYNFPQMAHQRFGRKRYQBLKEMR 246

QY 1358 WEREHQPDEDEDIKK 1374
Db 247 WEREHQPDEDEDIKK 263

RESULT 13
US-09-900-425A-3
; Sequence 3, Application US/09900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425A
```

```
/ CURRENT FILING DATE: 2002-01-29
/ PRIOR APPLICATION NUMBER: US 09/479,783
/ PRIOR FILING DATE: 2000-01-07
/ PRIOR APPLICATION NUMBER: US 08/870,608
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: US 80/659,440
/ PRIOR FILING DATE: 1996-06-06
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 3
/ LENGTH: 412
/ TYPE: PRT
/ ORGANISM: Caenorhabditis elegans
US-09-900-425A-3

Query Match      11.9%; Score 890; DB 3; Length 412;
Best Local Similarity 43.1%; Pred. No. 3e-46;
Matches 181; Conservative 85; Mismatches 136; Indels 18; Gaps 7;

QY 944 TLINIMSLGQDDPTPSRINHNRLERFLGDAVVEFLTSVHLYYLFPSLEGGGLATYRTAI 1003
Db 2 SLFNIMKGTSGGEP-----ILHNERLYLGDAVVELIVSHHLYFMLTHHFEGLATYRTAL 57

QY 1004 VQNHMLAKKLELDPFMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLEAKQL 1063
Db 58 VQNRNLATLAKNCRIDEMQLQYSHGADLINVAEFKHALANAFEAAMAIYLDGGLAPCDVI 117

QY 1064 FGRLLF-NDPDLREVWLNYPHLPLOQEPNTDQLIETSPVLOKLTFFEEAIGVIFTHVR 1122
Db 118 FSKAMYGHQPVLEKWDHINEHLKREDPQGDRLDSFITPTLSTFHAEERLGIQFNIR 177

QY 1123 LLARAFTLRTVGNHILTLGHNMREFLGSIMOLVATEYLFIHFPDHHGHLTLRLSSLV 1182
Db 118 FSKAMYGHQPVLEKWDHINEHLKREDPQGDRLDSFITPTLSTFHAEERLGIQFNIR 177

QY 1123 LLARAFTLRTVGNHILTLGHNMREFLGSIMOLVATEYLFIHFPDHHGHLTLRLSSLV 1182
Db 178 LLAKAFTRRNIPNNDLTGHNQRLKLEWGLSDVLQIVSDFLYRRFPYHHEGMSLLRTSLV 237

QY 1183 NNRTQAKVAEELGMOEYAITNDKTPV---GLRTKTLADLLESFIAALYTDKDLVYVHT 1239
Db 238 SNQTQAVVCDLGLTFEVI-----KAPYKTPPELKADKADLVEAFIGALYVDRGIEHCRA 292

QY 1240 FMNVCFPRLEKFIILNODWNPDSQLOQCCLTLR-TEGKEPDIPLYKTLQTVGPSHARTY 1298
Db 293 FIRIVECPRLKHFIESEKNDAKSHLQOWCLAMRDPSSSEPDMPYRVLGIEGPTNNRIF 352

QY 1299 TVAVYKGERIGCGKGPSIQOAEWGA---AMDALKYNNPQM-AHQKRFGRKYRQELKE 1354
Db 353 KIAVYKGRKLASAESNVHKAELRVAELALANLESMSFSKMAKNNSNRRLEQDTS 412

RESULT 14
US-10-079-185-3
/ Sequence 3, Application US/10079185
/ Publication No. US20030044941A1
/ GENERAL INFORMATION:
/ APPLICANT: Crooke, Stanley T.
/ TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
/ CURRENT APPLICATION NUMBER: US/10/079,185
/ PRIOR FILING DATE: 2002-02-20
/ PRIOR APPLICATION NUMBER: 09/479,783
/ PRIOR FILING DATE: 2000-01-07
/ PRIOR APPLICATION NUMBER: 08/870,608
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 08/659,440
/ PRIOR FILING DATE: 1996-06-06
/ PRIOR APPLICATION NUMBER: 09/900,425
/ PRIOR FILING DATE: 2001-07-06
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 3
/ LENGTH: 412
/ TYPE: PRT
/ ORGANISM: Caenorhabditis elegans
US-10-079-185-3

Query Match      11.9%; Score 890; DB 4; Length 412;
Best Local Similarity 43.1%; Pred. No. 3e-46;
Matches 181; Conservative 85; Mismatches 136; Indels 18; Gaps 7;

QY 944 TLINIMSLGQDDPTPSRINHNRLERFLGDAVVEFLTSVHLYYLFPSLEGGGLATYRTAI 1003
Db 2 SLFNIMKGTSGGEP-----ILHNERLYLGDAVVELIVSHHLYFMLTHHFEGLATYRTAL 57

QY 1004 VQNHMLAKKLELDPFMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLEAKQL 1063
Db 58 VQNRNLATLAKNCRIDEMQLQYSHGADLINVAEFKHALANAFEAAMAIYLDGGLAPCDVI 117

QY 1064 FGRLLF-NDPDLREVWLNYPHLPLOQEPNTDQLIETSPVLOKLTFFEEAIGVIFTHVR 1122
Db 118 FSKAMYGHQPVLEKWDHINEHLKREDPQGDRLDSFITPTLSTFHAEERLGIQFNIR 177

QY 1123 LLARAFTLRTVGNHILTLGHNMREFLGSIMOLVATEYLFIHFPDHHGHLTLRLSSLV 1182
Db 1183 LLARAFTLRTVGNHILTLGHNMREFLGSIMOLVATEYLFIHFPDHHGHLTLRLSSLV 1182
```


Db 178 LLAKAFTRRNIPNDLTKGHNQLEWLGDSVLQLIVSDFLYRRFPYHHEGHMSLLRTSLV 237
QY 1183 NNKTOAKVAEELGMOEYAITNDKTRPV---GLRTKTLADLLESFIAALYTDKLEYVHT 1239
Db 238 SNQTOAVVCCDLGTFEVI-----KAPYKTPPELKLKADLVEAFICALYVDRGIEHCRA 292
QY 1240 FNNVCFPRLEKEFILNQDNDPKSOLQOCCCLTLR-TEGKEPDIPLYKTLQTVGSPSHARTY 1298
Db 293 FIRIVFCPLKHFIESEKWNDAKSHLQQWCLAWRDPSSSEPDMPFYRVLGIEGFTNNRIF 352
QY 1299 TVAVYFGERIGCGKGPSIQQAEMGA---AMDAL EKYNFPQM-AHQKRFGRKYRQELKE 1354
Db 353 KIAVYKGRKLASAAESNVHKAELRVAELALANLESMSFSKMKAKNNSNNRRELEQDTS 412

Search completed: December 24, 2005, 01:02:56
Job time : 198 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 00:32:40 ; Search time 202 Seconds
(without alignments)
2988.646 Million cell updates/sec

Title: US-10-774-974-2
Perfect score: 7500
Sequence: 1 MQGNTCHMSFHPGRCPR.....MRWERHQREPDTEDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7500	100.0	1374	8 ADQ96675	Human rib
2	7500	100.0	1374	8 ADR68727	Human rib
3	7500	100.0	1374	9 AEB47411	Human rib
4	7486	99.8	1374	8 ADQ17464	Human sof
5	7486	99.8	1374	9 ADY17386	PRO polyp
6	7483.5	99.8	1373	7 ABU63361	Human dou
7	6426	85.7	1200	9 AEA20195	Novel hum
8	6062	80.8	1374	6 ABU08060	Human RNA
9	4071	54.3	769	4 AAB32635	Human pro
10	2963	39.5	1327	4 ABB58539	Drosophil
11	2847	38.0	541	6 ABM04841	Human put
12	2519	33.6	486	9 AEA21037	Novel hum
13	2463	32.8	466	8 ADQ96710	Human rib
14	1961	26.1	378	4 AAU20587	Human sec
15	1961	26.1	378	4 AAU21744	Novel hum
16	1961	26.1	378	7 ADC46385	Human neo
17	1956	26.1	378	4 AAU20385	Human sec
18	1320	17.6	263	4 AAU21580	Novel hum
19	1320	17.6	263	7 ADC46221	Human neo
20	1295.5	17.3	301	4 AAB63281	Human bre
21	1289	17.2	267	4 AAB63379	Human bre
22	1122	15.0	277	4 AAB63383	Human bre
23	890	11.9	412	8 ADQ96676	Caenorhab
24	890	11.9	412	8 ADR68728	C. elegans

25	890	11.9	412	9 AEB47412	Nematode
26	702	9.4	148	5 ADK34912	Novel hum
27	496	6.6	115	2 AAY12224	Human 5'
28	496	6.6	115	3 AAG00554	Human sec
29	472	6.3	97	5 ABB10067	Human non
30	340	4.5	1151	4 ABB61598	Abb10067 Drosophil
31	312.5	4.2	241	7 ADC95807	Adc95807 E. faeciu
32	311	4.1	560	4 ABG21040	Novel hum
33	309	4.1	286	9 AEB48256	Bacillus
34	307.5	4.1	229	5 ABB48891	Liasteria
35	303.5	4.0	406	4 ABG27250	Novel hum
36	301.5	4.0	1663	8 ABM81721	Abm81721 Tumour-as
37	298.5	4.0	228	5 ABP30268	Abp30268 Streptoco
38	298.5	4.0	228	8 ADV88449	Adv88449 Streptoco
39	298.5	4.0	228	8 ADV81861	Adv81861 Streptoco
40	298.5	4.0	228	8 ADV79702	Adv79702 Streptoco
41	298.5	4.0	242	5 ABP25995	Abp25995 Streptoco
42	297.5	4.0	800	8 ABO58564	AbO58564 Human gen
43	285.5	3.8	225	9 AEB48251	Vibrio ch
44	285.5	3.8	1527	8 ADU47047	Adu47047 Corn Dice
45	285.5	3.8	1636	8 ADU47051	Corn Dice

ALIGNMENTS

RESULT 1
ADQ96675
ID ADQ96675 standard; protein; 1374 AA.

XX AC ADQ96675;
XX XX
XX 23-SBP-2004 (first entry)
DT XX
DE DE Human ribonuclease III (RNase III) enzyme.

KW Human; ribonuclease III; RNase III; research purpose; biological purpose;
KW clinical purpose; cellular interaction; enzyme.
XX Homo sapiens.

Key Location/Qualifiers
Region 1..220 /note= "Proline rich region"
FT Misc-difference 18 /note= "Encoded by CGT"
FT Region 221..470 /note= "Serine-arginine rich region"
FT Domain 949..1374 /note= "RNase III domain"
FT Misc-difference 1211 /note= "Encoded by GCG"
FT Region 1262..1269 /note= "Alpha helix"
FT Region 1282..1290 /note= "Beta sheet"
FT Region 1297..1303 /note= "Beta sheet"
FT Region 1308..1315 /note= "Beta sheet"
FT Region 1316..1336 /note= "Alpha helix"
FT Misc-difference 1345 /note= "Encoded by GAA"

XX US2004126867-A1.
XX 01-JUL-2004.
XX 09-FEB-2004; 2004US-00774974.
XX 06-JUL-2001; 2001US-00900425.
XX

PA (CROO/) CROOKE S T.
XX (WUHH/) WU H.
PI Crooke ST, Wu H;
XX WPI; 2004-516913/49.
DR N-PSDB; ADQ96674.
XX
PT New isolated nucleic acid molecule encoding human RNase III, useful for
PT research, biological, or clinical purposes, e.g. defining the roles of
PT RNase III and the interaction of human RNase III and cellular RNA.
PS
XX Claim 3; SEQ ID NO 2; 31pp; English.
XX
CC The invention relates to human ribonuclease III (RNase III) and its
CC corresponding nucleic acid sequence. The polynucleotide sequence of the
CC invention is useful for research, biological and clinical purposes. It is
CC useful in defining the roles of RNase III and the interaction of human
CC RNase III and cellular RNA. The present sequence is the human RNase III
CC enzyme.
XX
SQ Sequence 1374 AA;
Query Match 100.0%; Score 7500; DB 8; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMQGNTRCHMSFHPGRCGRGCHGAPSPAPFRPNLRLLHPQQPPVQYQVEPPSAPS 60
DB 1 MMQGNTRCHMSFHPGRCGRGCHGAPSPAPFRPNLRLLHPQQPPVQYQVEPPSAPS 60
QY 61 TTSTNSPAPNPLPRDFVFPFPPMPSAQGLPPCEIRPPFNHQRHPPFPVPCFPDM 120
DB 61 TTSTNSPAPNPLPRDFVFPFPPMPSAQGLPPCEIRPPFNHQRHPPFPVPCFPDM 120
QY 121 PPMPCPNPVPVGPAGPGQTFPFMPMPSPMPPPPVMPQOVNTQYPPGYSHNFPPPP 180
DB 121 PPMPCPNPVPVGPAGPGQTFPFMPMPSPMPPPPVMPQOVNTQYPPGYSHNFPPPP 180
QY 181 SFNSFQNNPSSFLPSANSSSHPRLPPYPLPKAPSERRSERLKHDDHRDHSRGR 240
DB 181 SFNSFQNNPSSFLPSANSSSHPRLPPYPLPKAPSERRSERLKHDDHRDHSRGR 240
QY 241 GERHSLDRERGRSPDRRQDSRYSDYDGRGTPSHRSYERSRERERHHRNRRS 300
DB 241 GERHSLDRERGRSPDRRQDSRYSDYDGRGTPSHRSYERSRERERHHRNRRS 300
QY 301 PSLSRSYKKEYKRSYGLSVVPEPAGCTPELPGEIKNVDSWAPPLEIVNHRSPSREK 360
DB 301 PSLSRSYKKEYKRSYGLSVVPEPAGCTPELPGEIKNVDSWAPPLEIVNHRSPSREK 360
QY 361 KRARWEEKDRWSDNOSGKDKNYTSIKEKEPETPMDKNEEEEEELLLKPVWIRCTHSEN 420
DB 361 KRARWEEKDRWSDNOSGKDKNYTSIKEKEPETPMDKNEEEEEELLLKPVWIRCTHSEN 420
QY 421 YYSDDPMDQVGDSTVGTSLRLDYKFEELSGRQEKAKAAPPPWPPKTKLDELESS 480
DB 421 YYSDDPMDQVGDSTVGTSLRLDYKFEELSGRQEKAKAAPPPWPPKTKLDELESS 480
QY 481 SESECEDESDTSCSSSDSEVDVIAIEIKKKAHPDRHLDELWYNDPGQNDGPKLCKSA 540
DB 481 SESECEDESDTSCSSSDSEVDVIAIEIKKKAHPDRHLDELWYNDPGQNDGPKLCKSA 540
QY 541 KARRTGIRHSIYPGEEAIKFCRPMNNAAGLFHYRTVSPPTNFLTDRPTVIEYDDHEYI 600
DB 541 KARRTGIRHSIYPGEEAIKFCRPMNNAAGLFHYRTVSPPTNFLTDRPTVIEYDDHEYI 600
QY 601 FEGFSMAFAHAPLNIPLCKVIRFNIDYTHFIEEMMPENCVKGLELFSFLFRDILELY 660
DB 601 FEGFSMAFAHAPLNIPLCKVIRFNIDYTHFIEEMMPENCVKGLELFSFLFRDILELY 660
QY 661 DWNLKGPLFSDSPCCPRFHPMFRVFLPDGGKEVLSMHQILLYLLRCSKALVPEEEIA 720
DB 661 DWNLKGPLFSDSPCCPRFHPMFRVFLPDGGKEVLSMHQILLYLLRCSKALVPEEEIA 720

DB 661 DWNLKGPLFSDSPCCPRFHPMFRVFLPDGGKEVLSMHQILLYLLRCSKALVPEEEIA 720
QY 721 NMLQWEELEWQKYABECKGMIVTNPGTKPSSVRIQDLREQENPDVITPIIIVHFGIRPA 780
DB 721 NMLQWEELEWQKYABECKGMIVTNPGTKPSSVRIQDLREQENPDVITPIIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLWKSYYVKLRHLLANS PKVKTQDKQKLAQREBEALQKTRQKNTMRREVTV 840
DB 781 QLSYAGDPQYQKLWKSYYVKLRHLLANS PKVKTQDKQKLAQREBEALQKTRQKNTMRREVTV 840
QY 841 ELSQSGFWKTGIRSDVCQHMMPLVLTTHIRVHOCIMHLDKLGITYFQDRCCLQLAWTHP 900
DB 841 ELSQSGFWKTGIRSDVCQHMMPLVLTTHIRVHOCIMHLDKLGITYFQDRCCLQLAWTHP 900
QY 901 SHHLNFGNPNDRHARNSLNCGRQPKYGDVKVHHMHRKKGINTLINIMSRGLQDDPTPS 960
DB 901 SHHLNFGNPNDRHARNSLNCGRQPKYGDVKVHHMHRKKGINTLINIMSRGLQDDPTPS 960
QY 961 RINHNERLEFLGDVAVVEFLTSVHLVYLPFSLBEGGLATVYRTAIVQNQHMLAKKLELDP 1020
DB 961 RINHNERLEFLGDVAVVEFLTSVHLVYLPFSLBEGGLATVYRTAIVQNQHMLAKKLELDP 1020
QY 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEEAQKLFGRLLFNDPDLREVWLN 1080
DB 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEEAQKLFGRLLFNDPDLREVWLN 1080
QY 1081 YPLHPLQLOEPNTDROLIETSPVLOKLTEFEBAIGVIFTHVRLARAFRTLVGFNHLTL 1140
DB 1081 YPLHPLQLOEPNTDROLIETSPVLOKLTEFEBAIGVIFTHVRLARAFRTLVGFNHLTL 1140
QY 1141 GHNQRMFELGDSIMQIVATEYLFIFHPDHHGHTLLRSSLVNNRTQAKVABELGMOEYA 1200
DB 1141 GHNQRMFELGDSIMQIVATEYLFIFHPDHHGHTLLRSSLVNNRTQAKVABELGMOEYA 1200
QY 1201 ITNDTKRPVGLRTKTLADLLSFTAAALYTDKOLEVHTFMVAVCFPPRLKEFLINQDWD 1260
DB 1201 ITNDTKRPVGLRTKTLADLLSFTAAALYTDKOLEVHTFMVAVCFPPRLKEFLINQDWD 1260
QY 1261 PKSQOQCCCLTARTEGKPDIPLYKTQVGPSHARTYTVAVYFKGERIGCGKGPSIQQA 1320
DB 1261 PKSQOQCCCLTARTEGKPDIPLYKTQVGPSHARTYTVAVYFKGERIGCGKGPSIQQA 1320
QY 1321 EMGAAMDALKENFPQMAHQKRFGRYKQELKEMEWEREHQREPEDETDIKK 1374
DB 1321 EMGAAMDALKENFPQMAHQKRFGRYKQELKEMEWEREHQREPEDETDIKK 1374
RESULT 2
ADR68727
ID ADR68727 standard; protein; 1374 AA.
XX ADR68727;
DT 02-DEC-2004 (first entry)
XX Human Ribonuclease III, RNase III.
XX Human; enzyme; Ribonuclease III; RNase III; RNA interference;
KW gene silencing; double stranded RNA; pre-rRNA processing; RNA processing;
KW RNA expression; RNA splicing; RNA translocation.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 1. .220
Domain /label = Proline rich domain
FT /note= "Claimed in claim 8"
FT 221. .470
Domain /label = Serine Arginine rich domain
FT /note= "Claimed in claim 8"
FT 949. .1374
Domain /label = RNase III domain
FT /note= "Claimed in claim 8"

	Misc-difference 1211	/note= "Encoded by GCG"	
FT	Misc-difference 1345	/note= "Encoded by GAA"	
FT	US2004175828-A1.		
FT	09-SEP-2004.		
PN	22-MAR-2004; 2004US-00805919.		
XX	06-JUL-2001; 2001US-00900425.		
XX	(CROO/) CROOKE S T.		
PA	(WUHH/) WU H.		
XX	Crooke ST, Wu H;		
P1	WI; 2004-667111/65.		
XX	Modulating RNA interference, processing, expression, splicing and translocation in cell or tissue, by contacting cell or tissue with human RNase III or oligomeric compound targeted to nucleic acid encoding human RNase III.		
PT	Claim 7; SEQ ID NO 2; 32pp; English.		
PS	The invention relates to modulating RNA interference (gene silencing), RNA processing (e.g. of pre-rRNA), RNA expression, RNA splicing, or RNA translocation in a cell or tissue, by contacting the cell or tissue with a modulator effective to cause the modulation by at least 50% as compared to control, where modulator is human RNase III (ribonuclease III, which cleaves double stranded RNA) polypeptide or an oligomeric compound (an antisense oligonucleotide) targeted to the nucleic acid encoding human RNase III. The oligomeric compound is targeted to a 3'-untranslated region (3'UTR), a 5'-untranslated region (5'UTR) or a coding region of a nucleic acid molecule encoding human RNase III polypeptide, where the oligomeric compound inhibits the expression of human RNase III		
CC	polypeptide by at least 50%. In the method above, the RNA is rRNA, snRNA, snoRNA or miRNA, or precursors of the above. The RNA is processed to form one or more 30S and 32S RNA fragments. The 32S RNA is processed to form one or more 12S pre-rRNA and 28S rRNA fragments. The methods are useful for modulating RNA interference in a cell or tissue, modulating processing of an RNA in a cell or tissue, modulating RNA expression in a cell or tissue, modulating RNA splicing in a cell or tissue, and modulating RNA translocation in a cell or tissue. The methods are useful for research, biological and clinical purposes. The methods are useful in defining the roles of RNase III and interaction of human RNase III and cellular RNA. The present sequence represents human RNaseIII.		
CC	Sequence 1374 AA;		
SQ	Query Match 100.0%; Score 7500; DB 8; Length 1374; Best Local Similarity 100.0%; Pred. No. 0; Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MMGNTCHRMSPHPRGCRGGHGARGAPSPQNRLLLHPQQPVQYQYPPGPS 60		
DB	1 MMGNTCHRMSPHPRGCRGGHGARGAPSPQNRLLLHPQQPVQYQYPPGPS 60		
QY	61 TTFSNSPAPNPLPRPDFVFPFPPMPSPAQGLPFCPIRPFPFNHMRHPVPFPCFPPM 120		
DB	61 TTFSNSPAPNPLPRPDFVFPFPPMPSPAQGLPFCPIRPFPFNHMRHPVPFPCFPPM 120		
QY	121 PPMPCNNPPVPGAPGGCTFFMMPPSMHPHPPPPVMPQOVNYQYPGYSHHNPPP 180		
DB	121 PPMPCNNPPVPGAPGGCTFFMMPPSMHPHPPPPVMPQOVNYQYPGYSHHNPPP 180		
QY	181 SFNSFOFNTPSSFLPSANNSSPHFRHLPPYLPAKPSRRSRPERLKHYYDDHRHDHSHGR 240		
DB	181 SFNSFOFNTPSSFLPSANNSSPHFRHLPPYLPAKPSRRSRPERLKHYYDDHRHDHSHGR 240		
QY	241 GERHSILDRERGRSDRRQRDSRYRSYDGRGTFSRHRSYSRERERHRHRDNRRS 300		

Db	1321	EMGAAMDALEKYNFPQMAHQKRFGRKGRQELKEMWEREHQREPEDETEDIKK	1374
RESULT 3			
ID	AEBA7411		
XX	AEBA7411	standard; protein; 1374 AA.	
XX	AEBA7411;		
DT	22-SEP-2005	(first entry)	
XX			
DE	Human ribonuclease III (RNase III) protein, SEQ ID NO: 2.		
XX			
KW	Ribonuclease; gene silencing; enzyme.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Region	1..220	
FT		/note= "RNase III proline-rich region"	
FT	Misc-difference 18		
FT		/note= "Encoded by CGT"	
FT	Region	221..470	
FT		/note= "RNase III serine-arginine rich region"	
FT	Domain	949..1374	
FT		/note = RNase III domain	
FT	Misc-difference 1211		
FT		/note= "Encoded by GCG"	
FT	Region	1262..1269	
FT		/note= "Alpha helix"	
FT	Region	1282..1290	
FT		/note= "Beta sheet"	
FT	Region	1297..1303	
FT		/note= "Beta sheet"	
FT	Region	1308..1315	
FT		/note= "Beta sheet"	
FT	Region	1318..1336	
FT		/note= "Alpha helix"	
FT	Misc-difference 1345		
FT		/note= "Encoded by GAA"	
XX	US2005159384-A1.		
PN			
XX			
PD	21-JUL-2005.		
XX			
PF	02-DEC-2004; 2004US-00001993.		
XX			
PR	06-JUL-2001; 2001US-00900425.		
PR	20-FEB-2002; 2002US-00079185.		
XX	(ISIS-) ISIS PHARM INC.		
PA			
XX			
PI	Crooke ST;		
XX			
DR	WPI; 2005-512270/52.		
DR	N-PSDB; AEB47410.		
DR	GENBANK; AAF80558.		
XX			
PT	Use of RNase III for eliciting modification of a selected RNA target,		
PT	promoting gene silencing of a gene, inhibiting the expression of a gene,		
PT	promoting inhibition of expression of a gene, or eliciting modification		
PT	of an RNA target.		
XX			
PS	Claim 19; SEQ ID NO 2; 32pp; English.		
XX			
CC	The present invention relates to ribonuclease III (RNase III) nucleic		
CC	acids and their encoding proteins. RNase III is an endoribonuclease that		
CC	cleaves double stranded RNA. The invention is useful for eliciting		
CC	modification of a selected RNA target in a cell, promoting gene silencing		
CC	of a gene, inhibiting the expression of a gene, promoting inhibition of		
CC	expression of a gene and eliciting modification of a RNA target in a		
CC	cell. The present sequence is human ribonuclease III (RNase III) protein.		

SQ	Sequence	1374	AA;
	Query Match	100.0%;	Score 7500; DB 9; Length 1374;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 1374;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MMQGNCTCHRMSPHPCGRCGRGCGHGA	PSAPSFPQNRLRLHPQPPVQVQYEP
DB	1	MMQGNCTCHRMSPHPCGRCGRGCGHGA	PSAPSFPQNRLRLHPQPPVQVQYEP
QY	61	TTFSNSPAPNLFPPRDPFVPPPPMPPSA	QGLPCPIRPPPNHOMRPPVPPCFPPM
DB	61	TTFSNSPAPNLFPPRDPFVPPPPMPPSA	QGLPCPIRPPPNHOMRPPVPPCFPPM
QY	121	PPMPCPNPPVPAGAPPGQGTFFFMPPPS	MPHPPPPVMPQVNYQYPGYSHNFPPP
DB	121	PPMPCPNPPVPAGAPPGQGTFFFMPPPS	MPHPPPPVMPQVNYQYPGYSHNFPPP
QY	181	SFNSFQNNPSSFLPSANNSSSPHRLPYPL	PKAPSERRSERLKYDDHRRHDSHGR
DB	181	SFNSFQNNPSSFLPSANNSSSPHRLPYPL	PKAPSERRSERLKYDDHRRHDSHGR
QY	241	GERHSLSRRERGRSPDRRQDSRYRSYDR	GRGTSRHSYSRERERHRHRDRNR
DB	241	GERHSLSRRERGRSPDRRQDSRYRSYDR	GRGTSRHSYSRERERHRHRDRNR
QY	301	PSLSYKKEYKRSGRSYGLSVVPBPAGCT	PELPGEIIKNTDSWAPPLIIVNHRSP
DB	301	PSLSYKKEYKRSGRSYGLSVVPBPAGCT	PELPGEIIKNTDSWAPPLIIVNHRSP
QY	361	KKARWEEEEKDRWSDNQSSGDKNYTISI	KEKEPEETMPDKNEEEELLKPVWIR
DB	361	KKARWEEEEKDRWSDNQSSGDKNYTISI	KEKEPEETMPDKNEEEELLKPVWIR
QY	421	YYSSDPMDQVGDSTVVGTSRLRLDLDYD	KFBEELGSRQEKAKAAPPEPKTKLDE
DB	421	YYSSDPMDQVGDSTVVGTSRLRLDLDYD	KFBEELGSRQEKAKAAPPEPKTKLDE
QY	481	SESECESEDSTCSSSDSEVDFVIAEIKR	KAHPDRLHDELWYNDPGQWNDGPKCS
DB	481	SESECESEDSTCSSSDSEVDFVIAEIKR	KAHPDRLHDELWYNDPGQWNDGPKCS
QY	541	KARTGIRHSIYPGEAAIKCPMTNNAAGL	FHYRITVSPPTNFLTDRPTVIEYDDH
DB	541	KARTGIRHSIYPGEAAIKCPMTNNAAGL	FHYRITVSPPTNFLTDRPTVIEYDDH
QY	601	FEFGSMFAHAPLTNIPLCVKIRPNIDYTH	FTIEEMPPENFCVKGLELFSLFRDILE
DB	601	FEFGSMFAHAPLTNIPLCVKIRPNIDYTH	FTIEEMPPENFCVKGLELFSLFRDILE
QY	661	DNWLGKPLFEDSPCCPRFHEMPFRVRF	LPDGGKEVLSMHQIILLYLLRCSKALV
DB	661	DNWLGKPLFEDSPCCPRFHEMPFRVRF	LPDGGKEVLSMHQIILLYLLRCSKALV
QY	721	NMLQWEELEWQKYAECKGMIVTNPCTKP	SSVRAIDQLDREQFNPVITPPIIVHFG
DB	721	NMLQWEELEWQKYAECKGMIVTNPCTKP	SSVRAIDQLDREQFNPVITPPIIVHFG
QY	781	QLSYAGDPQVKLWKSIVYKLRHLLANS	PKVQTDKOKLAQREALOKIKRKNTR
DB	781	QLSYAGDPQVKLWKSIVYKLRHLLANS	PKVQTDKOKLAQREALOKIKRKNTR
QY	841	ELSSQGFVKTGIRSDYCOHAMMLPVLTH	IRVHQCLMHLDKLIGYTFQDRCLQLA
DB	841	ELSSQGFVKTGIRSDYCOHAMMLPVLTH	IRVHQCLMHLDKLIGYTFQDRCLQLA
QY	901	SHHLNFGMPDHARNLSNCGIRQPKYGR	KVHMHMRKKGINTLINMSRLGQDDTP
DB	901	SHHLNFGMPDHARNLSNCGIRQPKYGR	KVHMHMRKKGINTLINMSRLGQDDTP
QY	961	RINHNERLEFLGDAVVEFLTSVHLYYL	FPSSLEGGIATYTAIVQNOHLAMKLE
DB	961	RINHNERLEFLGDAVVEFLTSVHLYYL	FPSSLEGGIATYTAIVQNOHLAMKLE

QY 1021 FMLYAHGPDLCRESDLRHANCFEALIGAVYLEGSLEBAKOLFGRLLFNDPDLREYWLN 1080
DB 1021 FMLYAHGPDLCRESDLRHANCFEALIGAVYLEGSLEBAKOLFGRLLFNDPDLREYWLN 1080
QY 1081 YPLHPLQLQBPNTDROLIETSPVLQKLTPEEAIGVIFTHVRLLARAFTLRTVGFNHLTL 1140
DB 1081 YPLHPLQLQBPNTDROLIETSPVLQKLTPEEAIGVIFTHVRLLARAFTLRTVGFNHLTL 1140
QY 1141 GHNQMEFLGDSIMQLVATEYLFTHPPDHHEGLTLRLSSLVNVRTQAKVAEELGMOEYA 1200
DB 1141 GHNQMEFLGDSIMQLVATEYLFTHPPDHHEGLTLRLSSLVNVRTQAKVAEELGMOEYA 1200
QY 1201 ITNDKTRPVGLRTKTLADLLESFIAALYTDKOLEYVHTFMNVCFPPRLKEPILNQDWD 1260
DB 1201 ITNDKTRPVGLRTKTLADLLESFIAALYTDKOLEYVHTFMNVCFPPRLKEPILNQDWD 1260
QY 1261 PKSQLQCCCLTLTEGKEPDIPLYKTLQTVGPHSHARTYTVAVYFKGERICGKGPSIQQA 1320
DB 1261 PKSQLQCCCLTLTEGKEPDIPLYKTLQTVGPHSHARTYTVAVYFKGERICGKGPSIQQA 1320
QY 1321 EMGAAMDALKYFPQMAHQKFRIGRYROELKEMRWEREHQREPEDETIKK 1374
DB 1321 EMGAAMDALKYFPQMAHQKFRIGRYROELKEMRWEREHQREPEDETIKK 1374

RESULT 4

ADQ17464

ID ADQ17464 standard; protein; 1374 AA.

AC ADQ17464;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

26-AUG-2004 (first entry)

Human soft tissue sarcoma-upregulated protein - SEQ ID 281.

soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.

Homo sapiens.

WO2004048938-A2.

10-JUN-2004.

26-NOV-2003; 2003WO-US038193.

26-NOV-2002; 2002US-0429739P.

(PROT-) PROTEIN DESIGN LABS INC.

Aziz N, Ginsburg WM, Zlotnik A;

WPI; 2004-441208/41.

Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

Example 2; SEQ ID NO 281; 210pp; English.

The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytotstatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated protein of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.

XX

SQ

Sequence 1374 AA;

Query Match

Best Local Similarity

Matches 1372; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

QY

1

MMOQNTCHMSFHPGRCGRGCHGARPSAPFRPQNTLLHPQBPVQYQVEPPSAPS

60

DB

1

MMOQNTCHMSFHPGRCGRGCHGARPSAPFRPQNTLLHPQBPVQYQVEPPSAPS

60

QY

61

TTESNSPAPNLFPRDFVFPFPPMPPSACGLPPCEIRPFPNQHMRHFPVPPCCPPM

120

DB

61

TTESNSPAPNLFPRDFVFPFPPMPPSACGLPPCEIRPFPNQHMRHFPVPPCCPPM

120

QY

121

PPMPCCNPPVPVPCAPGGQTFPMMPPPPMPPPPPPVMPQVQVQYPPGYSHHNEPPP

180

DB

121

PPMPCCNPPVPVPCAPGGQTFPMMPPPPMPPPPPPVMPQVQVQYPPGYSHHNEPPP

180

QY

181

SFNSFQNTSSFLPSANSSSPHFRHLPPYPLPKAPSERRSPELKHYYDHRHDSHGR

240

DB

181

SFNSFQNTSSFLPSANSSSPHFRHLPPYPLPKAPSERRSPELKHYYDHRHDSHGR

240

QY

241

GERHSLDRRCGRSPDRRQDSRYSDYDRGTPSRHRSYERSRERERHRDRNRS

300

DB

241

GERHSLDRRCGRSPDRRQDSRYSDYDRGTPSRHRSYERSRERERHRDRNRS

300

QY

301

PSLSRYKKEYKRSGRSYGLSVVPEPAGCTPELPGIINKTDSWAPLEIVNHRSPREK

360

DB

301

PSLSRYKKEYKRSGRSYGLSVVPEPAGCTPELPGIINKTDSWAPLEIVNHRSPREK

360

QY

361

KXARWEEKDRMSDNQSSGDKNYTTSIXEKEPEETMPDKNEEEBELLKPVWIRCTHSEN

420

DB

361

KXARWEEKDRMSDNQSSGDKNYTTSIXEKEPEETMPDKNEEEBELLKPVWIRCTHSEN

420

QY

421

YSSDDPMDQVGDSTVVGTSRLRLDYKFEELGSRQEKAKAAPPPPEPKTKLDEDESS

480

DB

421

YSSDDPMDQVGDSTVVGTSRLRLDYKFEELGSRQEKAKAAPPPPEPKTKLDEDESS

480

QY

481

SESECESDESDTSCSSSDSEVPDVAEIKRKAHPDLRHLDELWYNDPGQNDGFLCKCSA

540

DB

481

SESECESDESDTSCSSSDSEVPDVAEIKRKAHPDLRHLDELWYNDPGQNDGFLCKCSA

540

QY

541

KARTGIRHSIYPGEEAIKPCRPMTNNAGRLFHYRITVSPPTNFLTDRPTVIEYDDHEYI

600

DB

541

KARTGIRHSIYPGEEAIKPCRPMTNNAGRLFHYRITVSPPTNFLTDRPTVIEYDDHEYI

600

QY

601

FEGFSNFAHAPLTNIPLCVKIRFNIDYTHFTBEEMPFNCVKGLFSLFLFRDILELY

660

DB

601

FEGFSNFAHAPLTNIPLCVKIRFNIDYTHFTBEEMPFNCVKGLFSLFLFRDILELY

660

QY

661

DWNLKGFLFEDSPCCPRFHEMPRPFVFLPDGSKVLSMHOILLYLLRCSKALVPBEEIA

720

DB

661

DWNLKGFLFEDSPCCPRFHEMPRPFVFLPDGSKVLSMHOILLYLLRCSKALVPBEEIA

720

QY

721

NMLQWEELEWQKYAECKGMIVTNPGCTKPSVRIQDLDRQFNPDVITPPIIVHFGIRPA

780

DB

721

NMLQWEELEWQKYAECKGMIVTNPGCTKPSVRIQDLDRQFNPDVITPPIIVHFGIRPA

780

QY

781

QLSYAGDPQYQKLWKSIVYKRLHLLANSPPKVTQDKQKLAQREBALQKIROKNTMRREVTV

840

DB

781

QLSYAGDPQYQKLWKSIVYKRLHLLANSPPKVTQDKQKLAQREBALQKIROKNTMRREVTV

840

QY

841

ELSSQCFWTKGRSDVCOHAMMLPVLTHIRHYHQCILMHLDKLIGYTFQDRCLLQAMTHP

900

DB

841

ELSSQCFWTKGRSDVCOHAMMLPVLTHIRHYHQCILMHLDKLIGYTFQDRCLLQAMTHP

900

QY

901

SHHLNFGMNPDPHARNSLSCGIRQPKYGRKVHMHMRKKGINTLINMSRLQDDPTPS

960

DB

901

SHHLNFGMNPDPHARNSLSCGIRQPKYGRKVHMHMRKKGINTLINMSRLQDDPTPS

960

QY

961

RINHNRBLFLGDVAVVEFITSVHLYLPPSLBEGGLATYRTAITVQNOHLAMLAUKLELDP

1020

||||| 661 RINHERLEFLGDAVVEFLTSVHLVYLFPSLEBGLATVYRTAIVQNOHLAMAKKLELDP 1020
1021 FMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLAEAKOLFGBLLFNDPDLREVWLN 1080
1021 FMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLAEAKOLFGBLLFNDPDLREVWLN 1080
1081 YPLHPLQLOQEPNTDROLIETSPVLQKLTFFEBAGVIFTHVRLARAFILRTVGFNHLTL 1140
1081 YPLHPLQLOQEPNTDROLIETSPVLQKLTFFEBAGVIFTHVRLARAFILRTVGFNHLTL 1140
1141 GHNQMEFLGDSIMQLVATEYLFIFHPDDHHEGLTLRLSSLVNRRTOAKVABEGLMOEYA 1200
1141 GHNQMEFLGDSIMQLVATEYLFIFHPDDHHEGLTLRLSSLVNRRTOAKVABEGLMOEYA 1200
1201 ITNDKTRPVGLTKTLADLLESFIAALYTDKOLEVHTFMNVCFPRPKLEFTLNQDWD 1260
1201 ITNDKTRPVGLTKTLADLLESFIAALYTDKOLEVHTFMNVCFPRPKLEFTLNQDWD 1260
1261 PKSOLOQCCLTLTTEGKEPDIPLYKTQVGPSHARTYTVAVYFKGERIGCGKGPSIQQA 1320
1261 PKSOLOQCCLTLTTEGKEPDIPLYKTQVGPSHARTYTVAVYFKGERIGCGKGPSIQQA 1320
1321 EMGAAMDALCKYNFPQMAHQKRIEGRKYRQELKEMWEREHQEREPEDEDIKK 1374
1321 EMGAAMDALCKYNFPQMAHQKRIEGRKYRQELKEMWEREHQEREPEDEDIKK 1374

RESULT 5
ADV17386
ID ADY17386 standard; protein; 1374 AA.
XX ADY17386;
DT
XX 05-MAY-2005 (first entry)
XX PRO polypeptide SEQ ID NO 3192.
DE
XX Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotrophic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; de; gene; diagnosis.
XX Homo sapiens.
OS
XX WO2005016962-A2.
PN
XX 24-FEB-2005.
PD
XX 11-AUG-2004; 2004WO-US026249.
XX
XX 11-AUG-2003; 2003US-0493546P.
XX (GETH) GENENTECH INC.
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX WPI; 2005-182330/19.
XX
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
XX Claim 8; SEQ ID NO 3192; 158pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a DNA encoding a PRO
CC polypeptide.
XX

SQ Sequence 1374 AA;
Query Match 99.8%; Score 7486; DB 9; Length 1374;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MMQGNCHRMSPHGRGCGRGGHGAARPSAPSPQNLRLHLPQPPVQVQYQYPPSAPS 60
DB 1 MMQGNCHRMSPHGRGCGRGGHGAARPSAPSPQNLRLHLPQPPVQVQYQYPPSAPS 60
QY 61 TTFSNSPAPNFPDPDFVPPPPMPPSAQAQGLPCCIRPPPPNQHMRHPPFVPPCFPPM 120
DB 61 TTFSNSPAPNFPDPDFVPPPPMPPSAQAQGLPCCIRPPPPNQHMRHPPFVPPCFPPM 120
QY 121 PPMPCPNPPVPVGPAGPPGQGTFFFMVPPPSMPPHPPPPVMPQOVNYYPPGYSHNFPFP 180
DB 121 PPMPCPNPPVPVGPAGPPGQGTFFFMVPPPSMPPHPPPPVMPQOVNYYPPGYSHNFPFP 180
QY 181 SFNSFQNNPSSFLPSANNSSPHFRLPPYPLPKAPSERRSERLKYDDHRRHDSHGR 240
DB 181 SFNSFQNNPSSFLPSANNSSPHFRLPPYPLPKAPSERRSERLKYDDHRRHDSHGR 240
QY 241 GERHSLDRERGRSPDRRRQDSRYSDYDRGTSPSRHSYERSRERERHRRDRNS 300
DB 241 GERHSLDRERGRSPDRRRQDSRYSDYDRGTSPSRHSYERSRERERHRRDRNS 300
QY 301 PSLSRYKKEYKRSGRSYGLSVVPPAGCTPBLPGEIIKNTDSWAPPLPVLNHRSPSRK 360
DB 301 PSLSRYKKEYKRSGRSYGLSVVPPAGCTPBLPGEIIKNTDSWAPPLPVLNHRSPSRK 360
QY 361 KRAWEEEXDRWSDNQSGKDKNYTISIKKEPEETMPDKNEEBEELLKPVWIRCTHSEN 420
DB 361 KRAWEEEXDRWSDNQSGKDKNYTISIKKEPEETMPDKNEEBEELLKPVWIRCTHSEN 420
QY 421 YVSSDPMDQVGSSTVGTSELRLDYDKFEEELGSRQEKAKAARPPWPKTKLDELESS 480
DB 421 YVSSDPMDQVGSSTVGTSELRLDYDKFEEELGSRQEKAKAARPPWPKTKLDELESS 480
QY 481 SESECESDSDTSCSSSDSEVDVIAEIKRKAHPDLRDLHDELWYNDPGQWMDGLCKCSA 540
DB 481 SESECESDSDTSCSSSDSEVDVIAEIKRKAHPDLRDLHDELWYNDPGQWMDGLCKCSA 540
QY 541 KARTGIRHSIYPGEAIAKPCPMTNNAQRLFHYRITVSPPTNFLTDRPTVYDDHEYI 600
DB 541 KARTGIRHSIYPGEAIAKPCPMTNNAQRLFHYRITVSPPTNFLTDRPTVYDDHEYI 600
QY 601 FEGFSMFAHAPLTNIDPLCKVIRPNIDYTHFTEEMMPENFCVKGLEFSLFLFRDILELY 660
DB 601 FEGFSMFAHAPLTNIDPLCKVIRPNIDYTHFTEEMMPENFCVKGLEFSLFLFRDILELY 660
QY 661 DWNKGLPFEDSPCCPRFHFMPRFVFLPDGKKEVLSMHQIILLYLLRCSKALVPBEEIA 720
DB 661 DWNKGLPFEDSPCCPRFHFMPRFVFLPDGKKEVLSMHQIILLYLLRCSKALVPBEEIA 720
QY 721 NMLQWEELEWQKYAECKGMIVTNPGTKSPSVRIDQDRQFNPDVITPPIVHFGIRPA 780
DB 721 NMLQWEELEWQKYAECKGMIVTNPGTKSPSVRIDQDRQFNPDVITPPIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLWKSIVKRLHLLANSKPVKQTDKOKLAQREALOKIRKNTMRREVTV 840
DB 781 QLSYAGDPQYQKLWKSIVKRLHLLANSKPVKQTDKOKLAQREALOKIRKNTMRREVTV 840
QY 841 ELSQSQGFWKTGIRSDVCOHAMMLPVLTTHIRYHQCLMHLDKLIGYTFQDRCLLQAMTHP 900
DB 841 ELSQSQGFWKTGIRSDVCOHAMMLPVLTTHIRYHQCLMHLDKLIGYTFQDRCLLQAMTHP 900
QY 901 SHLNFQMPNDHARNLSNCGIRQPKYGRKRVHMMNRKKGINTLINMSRLGDDTPPS 960
DB 901 SHLNFQMPNDHARNLSNCGIRQPKYGRKRVHMMNRKKGINTLINMSRLGDDTPPS 960
QY 961 RINHERLEFLGDAVVEFLTSVHLVYLFPSLEBGLATVYRTAIVQNOHLAMAKKLELDP 1020
DB 961 RINHERLEFLGDAVVEFLTSVHLVYLFPSLEBGLATVYRTAIVQNOHLAMAKKLELDP 1020

QY 1021 FMLYAHGPDLCRSDLRHANCFEALIGAVYLEGSLAEAKQLFGRLLFNDPDLREVWLN 1080
Db 1021 FMLYAHGPDLCRSDLRHANCFEALIGAVYLEGSLAEAKQLFGRLLFNDPDLREVWLN 1080
QY 1081 YPLHPLQLQBPNTDRQLIETSPVLQKLTEPEEAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
Db 1081 YPLHPLQLQBPNTDRQLIETSPVLQKLTEPEEAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
QY 1141 GHNORMEFLGDSIMQLVATYLFTHFPDHHGHLTLRLSSLVNRRTOAKVAEELGMQEYA 1200
Db 1141 GHNORMEFLGDSIMQLVATYLFTHFPDHHGHLTLRLSSLVNRRTOAKVAEELGMQEYA 1200
QY 1201 ITNDKTRPVGLTKTLADLLESFIAALYTDKLEYVHTFMVCFPPRLKEFTLNQDWD 1260
Db 1201 ITNDKTRPVGLTKTLADLLESFIAALYTDKLEYVHTFMVCFPPRLKEFTLNQDWD 1260
QY 1261 PKSQLOQCCLTLTEGKEPDIPLYKTLOTGPHSHARTYTVAVYFKGERICGKGPSIQQA 1320
Db 1261 PKSQLOQCCLTLTEGKEPDIPLYKTLOTGPHSHARTYTVAVYFKGERICGKGPSIQQA 1320
QY 1321 EMGAAMDALEKYNFPQMAHQKRFGRKRYQELKEMRWERHQBREPDETDIKK 1374
Db 1321 EMGAAMDALEKYNFPQMAHQKRFGRKRYQELKEMRWERHQBREPDETDIKK 1374

RESULT 6
ABU63361
ID ABU63361 standard; protein; 1373 AA.
AC ABU63361;
XX
DT 18-SEP-2003 (first entry)
DE Human double stranded RNase, RNase III.
XX
KW Human; enzyme; RNase III; double stranded RNase; RNA target;
KW gene silencing.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 18 /note= "Encoded by CGT"
FT Misc-difference 1211 /note= "Encoded by GCG"
FT Misc-difference 1345 /note= "Encoded by GAA"
XX
PN US2003044941-A1.
XX
PD 06-MAR-2003.
XX
PP 20-FEB-2002; 2002US-00079185.
XX
PR 06-JUN-1996; 96US-00659440.
PR 06-JUN-1997; 97US-00870608.
PR 07-JAN-2000; 2000US-00479783.
PR 06-JUL-2001; 2001US-00900425.
XX
PA (CROO/) CROOKE S T.
XX
PI Crooke ST;
XX
DR WPI; 2003-521756/49.
DR N-PSDB; ACD27526.
XX
PT Eliciting a modification of a selected RNA target in a cell, useful for
PT promoting inhibition of gene expression in a cell, comprises contacting
PT an RNA-like polynucleotide-RNA target duplex with a polypeptide having an
PT RNase III domain.
XX
PS Claim 37; Fig 1; 33pp; English.

XX The invention relates to eliciting a modification of a selected RNA
CC target in a cell comprises contacting an RNA-like polynucleotide-RNA
CC target duplex with a polypeptide comprising an RNase III domain. Also
CC included are promoting gene silencing in a cell, inhibiting the
CC expression of a gene in a cell comprising employing the method of cited
CC above, promoting inhibition of expression of a gene, a hybrid RNase III
CC (comprising at least one domain from a human RNase III and at least one
CC domain from an RNase III of an organism other than human) and a cell
CC having enhanced RNase III activity over an activity exhibited by a second
CC cell (where the second cell is not enriched with respect to the amount or
CC activity of RNase III polypeptide). The method is useful for eliciting a
CC modification of a selected RNA target in a cell, and for promoting
CC inhibition of expression of a gene in a cell. Compositions comprising
CC RNase III polypeptides or polynucleotides are useful for research,
CC biological and clinical purposes. The polynucleotides may be used in
CC defining the roles of RNase III and the interaction of human RNase III
CC and cellular RNA. Host cells can be used for the production of human
CC RNase III and for identifying agents, which increase or decrease levels
CC of expression or activity of human RNase III in the cell. The present
CC sequence represents human RNase III (a double stranded RNase)
XX
SQ Sequence 1373 AA;
Query Match 99.8%; Score 7483.5; D8 7; Length 1373;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1373; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MMQGNTRCHMSFHPGRGCGPRGCGHARGPSAPSRPNQLRLHLPQPPVQYQYEPSPAPS 60
Db 1 MMQGNTRCHMSFHPGRGCGPRGCGHARGPSAPSRPNQLRLHLPQPPVQYQYEPSPAPS 60
QY 61 TTFSNSPAPNLFPPDPFVFPFPPMPPSAQGLPPLPCPIRPPPHNQHRRHPPVPCFPMP 120
Db 61 TTFSNSPAPNLFPPDPFVFPFPPMPPSAQGLPPLPCPIRPPPHNQHRRHPPVPCFPMP 120
QY 121 PPMPCPNPPVPGAPPGGQTFPFMMPPSMHPHPPPPPPVMPQVQNTQYPGYSHHPPPP 180
Db 121 PPMPCPNPPVPGAPPGGQTFPFMMPPSMHPHPPPPPPVMPQVQNTQYPGYSHHPPPP 180
QY 181 SFNSFONNPSFLPSANNSSPHFRLPPYPLPKAPSERRSERLKHDDHRRDHSRGR 240
Db 181 SFNSFONNPSFLPSANNSSPHFRLPPYPLPKAPSERRSERLKHDDHRRDHSRGR 240
QY 241 GERHSLDRRGRSPDRRQDSRYSDYDRGTPSRHSYERSRERERHHRHRRNR 300
Db 241 GERHSLDRRGRSPDRRQDSRYSDYDRGTPSRHSYERSRERERHHRHRRNR 300
QY 301 PSLERSYKKEYKRSGRSYGLSVVPEPAGCTPBLPGBIIKNTDSWAPPLIIVNHRSPREK 360
Db 301 PSLERSYKKEYKRSGRSYGLSVVPEPAGCTPBLPGBIIKNTDSWAPPLIIVNHRSPREK 360
QY 361 KEARWEEKDRWSDNQSSGDKNYTISIKEKEPEETMPDKNEEEELLKPVVTRCTHSEN 420
Db 361 KEARWEEKDRWSDNQSSGDKNYTISIKEKEPEETMPDKNEEEELLKPVVTRCTHSEN 420
QY 421 YISSDPMDQVGDSTVVGTSRLRLDYDKFEEELGSRQEKAKAARPPWPKTKLDEDLSS 480
Db 421 YISSDPMDQVGDSTVVGTSRLRLDYDKFEEELGSRQEKAKAARPPWPKTKLDEDLSS 480
QY 481 SESECESDSDSTCSSSSDSEVDFVIAEIKRKAHPDRDLHDELWYNDPGQWNGPLCKCSA 540
Db 481 SESECESDSDSTCSSSSDSEVDFVIAEIKRKAHPDRDLHDELWYNDPGQWNGPLCKCSA 540
QY 541 KARRTGIRHSIYPGEAIPKCPMTNNAGRLPHYRTITVSPPTNFLTDRPTVLEYDDHEVI 600
Db 541 KARRTGIRHSIYPGEAIPKCPMTNNAGRLPHYRTITVSPPTNFLTDRPTVLEYDDHEVI 600
QY 601 FEGFSMFAHAPLTNIPCLCKVIRFNIDYTHFTEEMWPNFPCVKGLSFLFLFRDILELY 660
Db 601 FEGFSMFAHAPLTNIPCLCKVIRFNIDYTHFTEEMWPNFPCVKGLSFLFLFRDILELY 660
QY 661 DMNLKGPLPEDSPCCPRFHFMPRFVFLPDGCKEVLMSHQILLYLLRCSKALVPBEETA 720

Db 179 ----- 178

Qy 347 PLSIVNHRSPREKKEARWEEKDRWSDNQSGKDNYSIKKEPEETMPDKNEEEEE 406

Db 179 -----SRSPREKKEARWEEKDRWSDNQSGKDNYSIKKEPEETMPDKNEEEEE 232

Qy 407 LKPVWIRCTHSENYSSSDPMDQVGDSTVVGTSRLDLYDKFEEELGSRQEKAKAAPPW 466

Db 233 LKPVWIRCTHSENYSSSDPMDQVGDSTVVGTSRLDLYDKFEEELGSRQEKAKAAPPW 292

Qy 467 EPKTKLDELESSESESDSDSCSSSDSEVDVIAEIKRKAHPDRLHDELWYND 526

Db 293 EPKTKLDELESSESESDSDSCSSSDSEVDVIAEIKRKAHPDRLHDELWYND 352

Qy 527 PQQNDGPKCSAKAARTGIRHSIYPGEAIAKPCRPMTNNAAGRLPHYRITVSPPTNFLT 596

Db 353 PQQNDGPKCSAKAARTGIRHSIYPGEAIAKPCRPMTNNAAGRLPHYRITVSPPTNFLT 412

Qy 587 DRETVIYDDHEVIFPGFSFAHAPLNTNIPCLCKVIRFNIDYTHFTEEMPNFCVKGLE 646

Db 413 DRETVIYDDHEVIFPGFSFAHAPLNTNIPCLCKVIRFNIDYTHFTEEMPNFCVKGLE 472

Qy 647 LFSLFPRDLILEYDNLKGPLFEDGPPCCPRFHFPRFVRFLPDGCKEVLMSHQILLYL 706

Db 473 LFSLFPRDLILEYDNLKGPLFEDGPPCCPRFHFPRFVRFLPDGCKEVLMSHQILLYL 532

Qy 707 LRCSKALVPEEETIANLQWEELEWQYAECKGMIVTNPCTGTPSSVRIDQLDREQFNPDV 766

Db 533 LRCSKALVPEEETIANLQWEELEWQYAECKGMIVTNPCTGTPSSVRIDQLDREQFNPDV 592

Qy 767 ITPPIIVHFGIRPAQISYAGDPOYQKLWKSIVYKRLHLLANSPKVKQTDQKLAQREALQ 826

Db 593 ITPPIIVHFGIRPAQISYAGDPOYQKLWKSIVYKRLHLLANSPKVKQTDQKLAQREALQ 652

Qy 827 KIRQKNTMRREVTVELSSQGFMTGIRSDVCOHAMMLPVLTHIRYHQCIMHLDKLIGYT 886

Db 653 KIRQKNTMRREVTVELSSQGFMTGIRSDVCOHAMMLPVLTHIRYHQCIMHLDKLIGYT 712

Qy 887 FQDRCLLQALWTHPSHLNFGMNPDHARNSLSNCGIRQPKYQDRKVHMHMRKKGINTLI 946

Db 713 FQDRCLLQALWTHPSHLNFGMNPDHARNSLSNCGIRQPKYQDRKVHMHMRKKGINTLI 772

Qy 947 NIMSLRGQDDPTPSRINHNRLERFLGDVAVVEFTSVHLIYLPSPLEGGIATYRTAIVQN 1006

Db 773 NIMSLRGQDDPTPSRINHNRLERFLGDVAVVEFTSVHLIYLPSPLEGGIATYRTAIVQN 832

Qy 1007 QHLAMLAKKLELDPFMYAHGPDLCRESDLRHAWANCFEALIGAVYLEGSLREAKOLFGR 1066

Db 833 QHLAMLAKKLELDPFMYAHGPDLCRESDLRHAWANCFEALIGAVYLEGSLREAKOLFGR 892

Qy 1067 LFNPDPLREVWNLNYPHPLQLQEPNTDROLIETSPVLQKLTFFEBEAGIVFTHVRLAR 1126

Db 893 LFNPDPLREVWNLNYPHPLQLQEPNTDROLIETSPVLQKLTFFEBEAGIVFTHVRLAR 952

Qy 1127 AFTLRVTGPNHLLTGHNRQMEFLGDSIMQIVATEYLFPHPPDHHEGLTLLRSSLVNNT 1186

Db 953 AFTLRVTGPNHLLTGHNRQMEFLGDSIMQIVATEYLFPHPPDHHEGLTLLRSSLVNNT 1012

Qy 1187 QAKVAELGQVEYAITNDKTKRPGVTKTLADLLESFTAAALYTDKOLEVHTFMNVCF 1246

Db 1013 QAKVAELGQVEYAITNDKTKRPGVTKTLADLLESFTAAALYTDKOLEVHTFMNVCF 1072

Qy 1247 PRUKETILQDNDPKSQOCCCLTLRTEGKPDIPLYKTLQVGPSHARTYVAVYFKG 1306

Db 1073 PRUKETILQDNDPKSQOCCCLTLRTEGKPDIPLYKTLQVGPSHARTYVAVYFKG 1132

Qy 1307 ERIGCGKGPSIQAEAGAMDALEKYNFPQMAHQKRFGRKYRQELKEMRWEREHQEREP 1366

Db 1133 ERIGCGKGPSIQAEAGAMDALEKYNFPQMAHQKRFGRKYRQELKEMRWEREHQEREP 1192

Qy 1367 DETEDIKK 1374

|||||

Db 1193 DETEDIKK 1200

RESULT 8

ABU08060

ID ABU08060 standard; protein; 1374 AA.

XX

AC ABU08060;

XX

DT 16-MAY-2003 (first entry)

XX

Human RNase III protein.

XX

Human; enzyme; RNase III; ribonuclease III; endoribonuclease;

KW pre-ribosomal RNA; pre-rRNA; small molecular weight nuclear RNA; snRNA;

KW small molecular weight nuclear RNA; snRNA; mRNA degradation;

KW antisense therapy; RNA interference; RNAi; gene therapy;

XX

infectious agent; prophylaxis.

OS Homo sapiens.

XX

Key Location/Qualifiers

Region 161..168

FT /label= Alpha_helix_1

Region 181..189

FT /label= Beta_sheet_1

Region 196..202

FT /label= Beta_sheet_2

Region 207..214

FT /label= Beta_sheet_3

Region 215..235

FT /label= Alpha_helix_2

XX

US2002164601-A1.

PN

XX

07-NOV-2002.

PD

XX

06-JUL-2001; 2001US-00900425.

PF

XX

06-JUN-1996; 96US-00659440.

PR

06-JUN-1997; 97US-00870608.

PR

07-JAN-2000; 2000US-00479783.

XX

(WUHH/) WU H.

PA

(CROO/) CROOKE S T.

PA

XX

Wu H. Crooke ST;

PI

XX

WPI; 2003-328390/31.

DR

XX

New human RNase polypeptide, useful for screening antisense

PT oligonucleotides for therapy of disorders associated with RNase III

PT expression or activity, or for evaluating the efficacy of an antisense

PT therapy.

XX

Claim 3; Fig 1; 17pp; English.

PS

XX

The invention discloses an isolated human ribonuclease III (RNase III)

CC polypeptide, and the nucleic acid encoding it. RNase III is an

CC endoribonuclease that cleaves double stranded RNA. All RNase III species

CC contain an RNase III signature sequence. RNase III has been reported to

CC be involved in the processing of pre-ribosomal RNA (pre-rRNA), small

CC molecular weight nuclear RNAs (snRNAs) and small molecular weight

CC nuclear RNAs (snRNAs), as well as the degradation of some mRNA

CC species. Also disclosed is an antibody targeted to the human RNase III

CC polypeptide, an antisense compound 8 50 nucleobases in length, which is

CC targeted to the nucleic acid encoding human RNase III polypeptide, and

CC methods for inhibiting human RNase III expression, or activity, in a cell

CC or tissue, identifying agents that increase or decrease the activity or

CC levels of the human RNase III polypeptide in a host cell, screening

CC oligonucleotides to identify effective antisense oligonucleotides for

CC inhibition of expression of a selected target protein, prognosticating

CC efficacy of antisense therapy of a selected disease, eliciting cleavage

of a selected cellular RNA target and promoting RNA interference (RNAi) in a cell. The RNase III polypeptide, the polynucleotide encoding it and the antisense oligonucleotides, are useful for gene therapy (e.g. for treating a disease or disorder associated with RNase III expression or activity, or associated with an infectious agent), prophylaxis or as research reagents. The sequence presented is the human RNase III protein

Sequence 1374 AA;

Query Match 80.8%; Score 6062; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1101; Conservative 0; Mismatches 0; Indels 0

[illegible]

CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 769 AA;

Query Match 54.3%; Score 4071; DB 4; Length 769;
Best Local Similarity 99.5%; Pred. No. 4.2e-280;
Matches 765; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 606 MFAHAPLNTPLCKVIRFNIDYTHFTIEEMMPNFCVKGLELFLFRDILLEYDNLK 665
Db 1 MFAHAPLNTPLCKVIRFNIDYTHFTIEEMMPNFCVKGLELFLFRDILLEYDNLK 60
QY 666 GPLPEPSPCCPRFHFMPVRFLPDGGKEVLNMQHQLLYLLRCSKALVPEEIIANLQW 725
Db 61 GPLPEPSPCCPRFHFMPVRFLPDGGKEVLNMQHQLLYLLRCSKALVPEEIIANLQW 120
QY 726 EELEWKYAECKGMIVTNPSTPSSVRIDQLDREQFNPDVITPPIIVHFGIRPAQLSYA 785
Db 121 EELEWKYAECKGMIVTNPSTPSSVRIDQLDREQFNPDVITPPIIVHFGIRPAQLSYA 180
QY 786 GDPQYOKLWYSYKLRHLLANSKVKQTDKQKLAOREALQKIROKNTMRREVTVLSSQ 845
Db 181 GDPQYOKLWYSYKLRHLLANSKVKQTDKQKLAOREALQKIROKNTMRREVTVLSSQ 240
QY 846 GFWKTTGIRSDVCOHAMMLPVLTHIRVHOCLEHLDKLIQVTFODRCCLQLAMTHPSHLLN 905
Db 241 GFWKTTGIRSDVCOHAMMLPVLTHIRVHOCLEHLDKLIQVTFODRCCLQLAMTHPSHLLN 300
QY 906 FGNPDPHARNLSNCGIROPKYGDVKVHMHMKKGGINTLINMSRLGQDDPTPSRINHN 965
Db 301 FGNPDPHARNLSNCGIROPKYGDVKVHMHMKKGGINTLINMSRLGQDDPTPSRINHN 360
QY 966 ERLEFLGDAVVEFLTSVHLYLPPSLEEGGLATYRTAIVQNHMLAKKLELDPFMYA 1025
Db 361 ERLEFLGDAVVEFLTSVHLYLPPSLEEGGLATYRTAIVQNHMLAKKLELDPFMYA 420
QY 1026 HGPDLRESDLRHAMCNCFEALIGAVYLEGSLAEAKQLFGRLLFNDPDLREVLNPLHP 1085
Db 421 HGPDLRESDLRHAMCNCFEALIGAVYLEGSLAEAKQLFGRLLFNDPDLREVLNPLHP 480
QY 1086 LQLOEPTNDRLQLETSPVLQKLEFEEAIGVIFTHVRLARAFRTLVGFNHLTLGHNQR 1145
Db 481 LQLOEPTNDRLQLETSPVLQKLEFEEAIGVIFTHVRLARAFRTLVGFNHLTLGHNQR 540
QY 1146 MEFLGDSIMQLVATEYLFIHFPDHEGHLTLRLSSLVNNRTOAKVABELGMQSYAITNDK 1205
Db 541 MEFLGDSIMQLVATEYLFIHFPDHEGHLTLRLSSLVNNRTOAKVABELGMQSYAITNDK 600
QY 1206 TKRPVGLRTKTLADLLESFIAALYTDKLEYVTFMNVCFPRPKLEFILNQDNDPKSQL 1265
Db 601 TKRPVGLRTKTLADLLESFIAALYTDKLEYVTFMNVCFPRPKLEFILNQDNDPKSQL 660
QY 1266 QOCCLTIRTEGKPDPLPYKTLQTVGSHARTTVAVYFKGERIGCKGKPSIOQAEWGAA 1325
Db 661 QOCCLTIRTEGKPDPLPYKTLQTVGSHARTTVAVYFKGERIGCKGKPSIOQAEWGAA 720
QY 1326 MDALKYNFPQMAHQKRFIRGRKYQELKEMRWEREHQERPDTEDIKK 1374
Db 721 MDALKYNFPQMAHQKRFIRGRKYQELKEMRWEREHQERPDTEDIKK 769

RESULT 10
ABB58539
ID ABB58539 standard; protein; 1327 AA.
XX
AC ABB58539;
XX
DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 2409.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US009231.
PP
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
PR
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
XX N-PSDB; ABL02642.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
PT
XX Disclosure; SEQ ID NO 2409; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1327 AA;

Query Match 39.5%; Score 2963; DB 4; Length 1327;
Best Local Similarity 45.6%; Pred. No. 4.7e-201;
Matches 621; Conservative 213; Mismatches 327; Indels 200; Gaps 32;

QY 90 QGPLPCPIRPPPPNHQMRHPPVPCFPMPMPMPCPN-NPPVPGAP-----P 137
Db 3 QPELPPPPVQPA-----PPPPPPPEEDLSPPGVGVPSHNYSSNESHQ 46
QY 138 GQGTFFPMPPSPMPPPPPPVMPVMPQVNYQYPPGY-----SHNFP 179
Db 47 SSKSLDYVVPETAPYASS---VPSYDPYQOPPAYGYEGVAYNEQAQYGGQSHYQY 103
QY 180 PSFNSQNNPSSFLPSANNSSPHRHLPPYPL-----PKAPSRRRSPERLKHVDHHRD 235
Db 104 PA-----SGSSFLYES-----YKYPDRYPAYSSNYRPPSERQ-----RYTS 139
QY 236 HSHGRGERHSLDRREGSPDRRRQDSYRSYDRGRTPSRHSYRSEREREHRR 295
Db 140 NSSSQYHYHP-----GYSSGR-----RYEORHQ-----EHRQODSYAHPRHGHY 183
QY 296 DNRSPSLERSYKKEYKRSRSGYGLSWPEPAGCTPELPGCEIKNTDSSWAPPLEIVNHR 355
Db 184 AHRQAKGSHGY---YGSARN-----QVSDDYSPRGHERN 219
QY 356 PSREKRA--RWEEKDR-----WSDNQSGKOKNNTS1KEKSEETMPDKNEEBELK 409
Db 220 ETLEKTRAKPKVETERDLRLRQWCSNFC-----EKPEDYVKKNALSEADAPV 267
QY 410 FWIRCTSHENYSSDFMDQVGDSTVGTSLRLDLTKFEELSGRQEKAKARPPEPP 469

KW clinical purpose; cellular interaction.
 XX Homo sapiens.
 OS US2004126867-A1.
 PN XX
 XX 01-JUL-2004.
 PD XX
 XX 09-FEB-2004; 2004US-00774974.
 PF XX
 XX 06-JUL-2001; 2001US-00900425.
 PR XX
 XX (CROO/) CROOKE S T.
 PA (WUHH/) WU H.
 PA XX
 XX Crooke ST, Wu H;
 PI XX
 XX WPI; 2004-516913/49.
 DR XX
 XX New isolated nucleic acid molecule encoding human RNase III, useful for
 PT research, biological, or clinical purposes, e.g. defining the roles of
 PT RNase III and the interaction of human RNase III and cellular RNA.
 XX
 XX Example 9; SEQ ID NO 37; 31bp; English.
 PS XX
 XX The invention relates to human ribonuclease III (RNase III) and its
 CC corresponding nucleic acid sequence. The polynucleotide sequence of the
 CC invention is useful for research, biological and clinical purposes. It is
 CC useful in defining the roles of RNase III and the interaction of human
 CC RNase III and cellular RNA. The present sequence is human RNase III-like
 CC domain protein.
 CC
 XX
 SQ Sequence 466 AA;
 Query Match 32.8%; Score 2463; DB 8; Length 466;
 Best Local Similarity 100.0%; Pred. No. 4e-166;
 Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 909 NPDHARSLNSGCIROPKYGDVKVHHMMKKGKINTLIMSLRGQDDPTPSRININERL 968
 Db 1 NPDHARSLNSGCIROPKYGDVKVHHMMKKGKINTLIMSLRGQDDPTPSRININERL 60
 QY 969 EFLGDVAVVEFTSVHLYLFPSEEGGLATYRTAIVQNHLMAMAKKLEDPFMYAHGP 1028
 Db 61 EFLGDVAVVEFTSVHLYLFPSEEGGLATYRTAIVQNHLMAMAKKLEDPFMYAHGP 120
 QY 1029 DLCRESDLHAMANCFEALIGAVYLEGSLEAKQLFGRLLFNDPDLREVWLNYPHLPLQL 1088
 Db 121 DLCRESDLHAMANCFEALIGAVYLEGSLEAKQLFGRLLFNDPDLREVWLNYPHLPLQL 180
 QY 1089 QEPNDRQLIETSPVLOKLTFEPEAIGVIFTHVRLARAFPLRTVGFNHLTLGHNORMEP 1148
 Db 181 QEPNDRQLIETSPVLOKLTFEPEAIGVIFTHVRLARAFPLRTVGFNHLTLGHNORMEP 240
 QY 1149 LGSIMQLVATEYLFIHFPDPHHEGLTLRLSSLVNNTQAKVAEELGMQSYAITNDKTKR 1208
 Db 241 LGSIMQLVATEYLFIHFPDPHHEGLTLRLSSLVNNTQAKVAEELGMQSYAITNDKTKR 300
 QY 1209 PVGLRTKTLADLLESFTAAALYTDKLEYVHTFMNVCFFPRLKEFILNQDNDPKSQLOQC 1268
 Db 301 PVGLRTKTLADLLESFTAAALYTDKLEYVHTFMNVCFFPRLKEFILNQDNDPKSQLOQC 360
 QY 1269 CLTLRTGKEPDIPLYKTLQTVGSHARTYTVAVYFGERIGCGKGPSIQQAEMGAAMD 1328
 Db 361 CLTLRTGKEPDIPLYKTLQTVGSHARTYTVAVYFGERIGCGKGPSIQQAEMGAAMD 420
 QY 1329 LEKNYFQMAHQRFICRKYRQELKEMRWEREHQEREPEDETIKK 1374
 Db 421 LEKNYFQMAHQRFICRKYRQELKEMRWEREHQEREPEDETIKK 466

Db 181 LRSSLVNRTOAKVAEELGMEQVAITNDKTRPVALRTKTLADLLRSFIAALYIDKDLEY 240
Qy 1237 VHTFMNVCFPRRIKEPILNQDNDPKSLOQCCILTRTEGKEPDIPLYKTLOTVGPSHAR 1296
Db 241 VHTFMNVCFPRRIKEPILNQDNDPKSLOQCCILTRTEGKEPDIPLYKTLOTVGPSHAR 300
Qy 1297 TYTVAVYFGERIGCGKGPSIQAEAGAAADALEKYNFQMAHQKRFIGRKYQELKEWR 1356
Db 301 TYTVAVYFGERIGCGKGPSIQAEAGAAADALEKYNFQMAHQKRFIERKYQELKEWR 360
Qy 1357 WEREHQEREDETEDIKK 1374
Db 361 WEREHQEREDETEDIKK 378
RESULT 15
AAU211744
ID AAU211744 standard; protein; 378 AA.
AC AAU211744;
XX
DT 04-DEC-2001 (first entry)
XX
DE Novel human neoplastic disease associated polypeptide #177.
XX
KW Human; neoplastic disease associated polypeptide; cancer;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasotropic.
XX
OS Homo sapiens.
XX
PN WO20015163-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001358.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 28-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0233223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239353P.
PR 13-OCT-2000; 2000US-023937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.

PR	17-NOV-2000;	2000US-0249210P.	
PR	17-NOV-2000;	2000US-0249211P.	
PR	17-NOV-2000;	2000US-0249212P.	
PR	17-NOV-2000;	2000US-0249213P.	
PR	17-NOV-2000;	2000US-0249214P.	
PR	17-NOV-2000;	2000US-0249215P.	
PR	17-NOV-2000;	2000US-0249216P.	
PR	17-NOV-2000;	2000US-0249217P.	
PR	17-NOV-2000;	2000US-0249218P.	
PR	17-NOV-2000;	2000US-0249244P.	
PR	17-NOV-2000;	2000US-0249245P.	
PR	17-NOV-2000;	2000US-0249264P.	
PR	17-NOV-2000;	2000US-0249265P.	
PR	17-NOV-2000;	2000US-0249297P.	
PR	17-NOV-2000;	2000US-0249299P.	
PR	17-NOV-2000;	2000US-0249300P.	
PR	01-DEC-2000;	2000US-0250160P.	
PR	01-DEC-2000;	2000US-0250391P.	
PR	05-DEC-2000;	2000US-0251030P.	
PR	05-DEC-2000;	2000US-0251988P.	
PR	06-DEC-2000;	2000US-0256719P.	
PR	06-DEC-2000;	2000US-0251479P.	
PR	08-DEC-2000;	2000US-0251856P.	
PR	08-DEC-2000;	2000US-0251868P.	
PR	08-DEC-2000;	2000US-0251869P.	
PR	08-DEC-2000;	2000US-0251989P.	
PR	08-DEC-2000;	2000US-0251990P.	
PR	11-DEC-2000;	2000US-0254097P.	
PR	05-JAN-2001;	2001US-0259678P.	
XX	(HUMA-) HUMAN GENOME SCI INC.		
XX			
XX	Rosen CA, Barash SC, Ruben SM;		
XX	WPI; 2001-465558/50.		
DR	N-PSDB; AAS34943.		
XX			
PT	Novel polypeptides and polynucleotides useful as diagnostic reagents to		
PT	diagnose diseases or disorders associated with aberrant expression or		
PT	activity of polypeptides, and for treating cancers, rheumatoid arthritis.		
XX			
PS	Claim 11; SEQ ID NO 471; 687pp; English.		
XX			
CC	The present invention relates to the isolation of novel human neoplastic		
CC	disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA		
CC	sequences encoding for these polypeptides. The sequences of the invention		
CC	are useful in the diagnosis, treatment, prevention and/or prognosis of		
CC	disorders involving neoplastic disease such as hyperproliferative		
CC	disorders (e.g. leukaemia, bone cancer, bladder cancer, brain stem		
CC	glioma, adult liver cancer, childhood cerebellar astrocytoma, or		
CC	Hodgkin's lymphoma). The sequences of the invention may also be useful		
CC	for treating other disorders such as neural disorders, immune system		
CC	disorders, muscular disorders, reproductive disorders, gastrointestinal		
CC	disorders, pulmonary disorders, cardiovascular disorders and renal		
CC	disorders. The polynucleotide sequences of the invention are also useful		
CC	in gene therapy. AAU21568-AAU21851 represent the novel human neoplastic		
CC	disease associated polypeptides of the invention. Note: The sequence data		
CC	for this patent did not form part of the printed specification, but was		
CC	obtained in electronic format directly from WIPO at		
CC	ftp.wipo.int/pub/published_pct_sequences		
XX			
XX	Sequence 378 AA;		

Query Match		26.1%	Score 1961;	DB 4;	Length 378;
Best Local Similarity		98.9%	Pred. No. 1.4e-130;		
Matches 374;		Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	997	ATYRTAIVQNHQHLAMLAKKLELDPFMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGS	1056		
DB	1	ATYRTAIVQNHQHLAMLAKKLELDRFMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGS	60		
QY	1057	LEEAQOLFGRLLFNDPDLREVWLNYPHLPLQLOEPNTDRLIETSPVLQKLTFFEEAIGV	1116		

Db	61	LEEAQOLFGRLLFNDPDLREVWLNYPHLPLQLOEPNTDRLIETSPVLQKLTFFEEAIGV	120
QY	1117	IFTHVRLARAFPLRTVGFNHLTLGHNQRMFGLDSIMQLVATEYLFIHPPDHHEGHLTL	1176
Db	121	IFTHVRLARAFPLRTVGFNHLTLGHNQRMFGLDSIMQLVATEYLFIHPPDHHEGHLTL	180
QY	1177	LRSSLVNNRTQAKVAEELGMQEVATINDKTKRPVGLRTKTLADLLESFIAALYTDKLEY	1236
Db	181	LRSSLVNNRTQAKVAEELGMQEVATINDKTKRPVGLRTKTLADLLESFIAALYTDKLEY	240
QY	1237	VHTFMNVCFPPRLKEFILNQDNDPKSQLOQCCLTLRTGKESPDIPLYKTLQTVGFSHAR	1296
Db	241	VHTFMNVCFPPRLKEFILNQDNDPKSQLOQCCLTLRTGKESPDIPLYKTLQTVGFSHAR	300
QY	1297	TYTVAVYFKGERIGCGKGPSIQQAENGAAMDALAEKTNFPQMAHQKRPICRKYQELKEMR	1356
Db	301	TYTVAVYFKGERIGCGKGPSIQQAENGAAMDALAEKTNFPQMAHQKRPICRKYQELKEMR	360
QY	1357	WEREHQEREPEDETIKK	1374
Db	361	WEREHQEREPEDETIKK	378

Search completed: December 24, 2005, 00:53:40
Job time : 212 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 00:35:55 ; Search time 29 Seconds
(without alignments)
4558.684 Million cell updates/sec

Title: US-10-774-974-2
Perfect score: 7500
Sequence: 1 MMQGYTCHRMSPHFGGCRP.....MRWERHQERPEDETIKK 1374
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	11.9	412	2 T21419	hypothetical prote
2	558	7.4	682	2 T21420	hypothetical prote
3	316	4.2	551	2 S57447	HPBRII-7 protein -
4	315	4.2	249	2 B69693	ribonuclease III (
5	307.5	4.1	229	2 A81300	ribonuclease III h
6	300	4.0	229	2 A81672	ribonuclease III h
7	293	3.9	1560	2 T42727	proliferation pote
8	285.5	3.8	225	2 B82073	ribonuclease III v
9	282	3.8	263	2 A83961	ribonuclease III x
10	281.5	3.8	232	2 P98012	ribonuclease III (
11	279.5	3.7	232	2 H95144	ribonuclease III (
12	276	3.7	948	2 A57640	retinoblastoma bin
13	275	3.7	760	2 T06291	extensin homolog T
14	274.5	3.7	1531	2 T48946	hypothetical prote
15	273.5	3.6	1006	2 G86292	hypothetical prote
16	273	3.6	891	2 G84693	probable proline-r
17	271.5	3.6	620	2 S06733	hydroxyproline-ric
18	270	3.6	226	2 B84959	ribonuclease III (
19	268.5	3.6	231	2 H86725	ribonuclease III (
20	267	3.6	224	2 A81260	ribonuclease III (
21	267	3.6	243	2 D89896	RNase III (impor
22	267	3.6	998	2 T30930	hypothetical prote
23	265.5	3.5	1002	2 T43236	carboxypeptidase C
24	262	3.5	226	1 NREC3	ribonuclease III (
25	262	3.5	226	2 F85902	RNase III, ds RNA
26	262	3.5	226	2 A91058	RNase III (impor
27	259	3.5	322	2 S25299	extensin precursor
28	258	3.4	231	2 S76204	hypothetical prote
29	257.5	3.4	272	2 T35656	ribonuclease III -

30	256	3.4	246	2 H70187	ribonuclease III (
31	255.5	3.4	1638	2 A42091	transcription acti
32	255	3.4	226	2 AC0829	ribonuclease III (
33	254	3.4	1611	2 T38236	hypothetical prote
34	253.5	3.4	368	2 C29356	hydroxyproline-ric
35	253.5	3.4	1048	2 T31425	C-terminal domain-
36	252	3.4	226	2 AP0331	ribonuclease III (
37	251.5	3.4	240	2 P64602	ribonuclease III -
38	251.5	3.4	1201	2 G86441	unknown protein (i
39	250.5	3.3	727	2 C84534	hypothetical prote
40	249	3.3	225	2 T09964	extensin CYC15 pre
41	249	3.3	907	2 E96636	hypothetical prote
42	249	3.3	1877	2 T21861	hypothetical prote
43	245	3.3	230	2 G97115	dsRNA-specific rib
44	244.5	3.3	489	2 T11622	extensin class 1 p
45	243.5	3.2	708	2 D96711	hypothetical prote

ALIGNMENTS

RESULT 1

T21419
hypothetical protein F26E4.13 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21419
R:Lightning, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19419
A:Accession: T21419
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-412 <WIL>
A:Cross-references: UNIPARC:UPI000017B936; EMBL:Z81070; PIDN:Z81070; GSPDB:GN000019; (;
A:Experimental source: clone F26E4
C:Genetics:
A:Gene: CESP:F26E4.13
A:Map position: 1
A:Introns: 359/2; 400/3

Query Match		11.9%	Score 890;	DB 2;	Length 412;
Best Local Similarity		43.1%;	Pred. No. 6.6e-40;		
Matches 181;		Conservative 85;	Mismatches 136;	Indels 18;	Gaps 7;
QY	944	TLINIMSRIGQDDPTSPRINHNRELFGLDGVAVFLTSVHLVYLPSPLEGGATYRTAI	1003		
DB	2	SLFNINKGTSGGEP-----ILHNERLEVLDGAVVELIVSHLYFMLTHHFGGLATYRTAL	57		
QY	1004	VQNHMLAKKLELDPFMYAHGPDLCHESDLRHMANCFEALIGAVYLEGSLERAKQL	1063		
DB	58	VQNRNLATLAKNCRIDEMLQYSHGADLINVAEPKHALANAFEAVMAAIYLDGGLAPCDVI	117		
QY	1064	FGRLLF-NDPDLREVWLYPLHPLQEQPNTDRLQIETSPVLQKLTFFREAIQVIFTHVR	1122		
DB	118	FSKAMYGHOPVLKPKRWHDHNEHKLKEDPQGDRLDSFIPTTLSTFHAEERLGIQFNIR	177		
QY	1123	LLARAFTLTVTFNHLTLGHNRMEFLGDSIMOLVATEYLFIHFDPDHEGHLTLKSLV	1182		
DB	178	LLAKATRRNIPNNDLTKGHNQLEWLSVLQIIVSDFLYRFPFYHHEGHSLLRTSLV	237		
QY	1183	NNRTQAKVAEELGMQBYAITNDKTRPV---GLRTKTLADLLSFTAAALYTDKLEVHT	1239		
DB	238	SNQTAQVVDLLGTFTEFVI-----KAPYKTPPELKDADLVEAFICALYVDRIEGHCRA	292		
QY	1240	FMNVCFPLKKEPILNQDNDPKSQLOQCCLTLR-TEGKEPDIPLYKTLQTVGPSHARTY	1298		
DB	293	FRIVFCPLKXHFIESEKWNDAKSHLQOCLAMRWDPSSSEPDMPFVRVLGIEGTNNRIF	352		
QY	1299	TVAVYFKGRIGCGKGPSIQQAEMGA---AMDALKYFNPM-AHQKRFITGRKYRBLKE	1354		
DB	353	KIATVYKGRILASAESNVHKAELRVLAELALANLESMSFSKMKAKNNSNRRLEQDTS	412		

```
RESULT 2
T21420
hypothetical protein F26E4.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21420
R:Lightning, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19419
A:Accession: T21420
A>Status: preliminary; translated from GB/EMBL/DDJ
A:Molecule type: DNA
A:Residues: 1-682 <WIL>
A:Cross-references: UNIPARC:UPI0000179272; EMBL:Z81070; PIDN:CAB03006.1; GSPDB:GN000019;
A:Experimental source: clone F26E4
C:Genetics:
A:Gene: CESP:F26E4.10
A:Map position: 1
A:Introns: 119/3; 226/2; 303/2; 377/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F26E4.10

Query Match          7.4%; Score 558; DB 2; Length 682;
Best Local Similarity 24.1%; Pred. No. 4.4e-22;
Matches 178; Conservative 126; Mismatches 296; Indels 140; Gaps 23;

QY 277 RHRSYRSRERERHRHRDNRRSPSLERSYKKEYKSGSLGVVPEPAGCTPELP-G 335
DB 14 KHKRARKKYQKEYQERHKEEMMOQLGRRPQNPSTSSAPDPVVEKIPLTSTSLPFG 73

QY 336 EIIKNTDS-----WAPLLEIVNHRSPREKKRARWEBEKRDWSNQSGDKNVTSIK 388
DB 74 DSPRLTEKQYETNYMDPP--VVVTHSAELIK-----SNRWVVK 110

QY 389 EKEPEPTMDKNSEEBEELK-----PVMIRCTHS----- 418
DB 111 ABEAKYMIKAKSTTKILQDPQTKILETVTKRRLQADVPYIIHPCHSMGKRKTPKQK 170

QY 419 ---ENYSSDPMDOVGDSVTWGTSLRLDLYKFEELGSGEQEAKAARPP--WEPPKTKL 473
DB 171 GGDESTASDVSDSDS-----QDEASTSEPTNRQAPSAKDT 208

QY 474 DEDLESSESECSDDSTCS-----SSSDSEVDVTAIEIKRKAHPDRLH 519
DB 209 GE-----VKDEKQICNRRNQKAKRLNFEKEKQIITLLKKIDRKKTHNGIH 258

QY 520 DELWYNDPGMNDPGICKSAKARTGIRHSIIVGSEAIKPCRPMTNNAGRLPHY--RIT 577
DB 259 PDISFNEKGLNEGPEKCPPIKTCGLKHGYVAGEDKAIDCK--KXNGENLHYITLRTV 316

QY 578 VSPPTNFLTDRPTVIEYDDHEIYFEGFSMPAHAPL-----TNIPLCVKVIRENIDYTHFIE 633
DB 317 PLPSENQL--YRTHMAINGEFEGFSLTHAPLPDCTRAPICK---YSMDYEFOLVE 371

QY 634 EMMP-ENFCVKGLFSLFDRILEYLDMNLGPLFEDSPCCPRFHFMPRVFRLPDG 692
DB 372 EFWPDECFDEDCMLFEYIFHEIFEMLDPELAPKHPSPVESCPMIHMPRVQ--TKDD 430

QY 693 GKEVLMSHQILLYLL--RCSKALVPEEIANMLQWELEWOKYABECKGMIVTNPQKPS 751
DB 431 LVQLWSSKTVLAVFTSGSSEIMSPEDVNLCDQIDQFTRNTSKHKQSVILNFKFPSA 490

QY 752 VRIDQLDREQNPDVTFPIIVHFGIRPAQLSVAGDPQYOKLWKSYYKLRHLLANSKPKYK 811
DB 491 IRADWFERDEKEKEV-----YVFNHNAIRAQTYTATSLPRIAFLEKTLNKMIQ--EKQSSGVY 545

QY 812 QTDOKLAQREAEALOKIRKNTWREVTVELSQGWKTKGIRSDVCOHAMMLPVLTHHIR 871
DB 546 NKDFEK---TKNELEHLKRENRSARNKLREPVAGFTETGLKPDVAHVVTWTLACHHIR 602

QY 872 YHQCMLHDLKLIYTTQDRCLLQAMTHPSHLLNFGMNPDPHARNSLSNCGIRPKYGD RK 931
DB 603 YNFSLDVFEVIEYKFNDRRVIELALMHSSFKSHYGTPIDHVKNMITNCCGYR--KYG--- 658
```

```
QY 932 VHHMHRKKGINTLINISR 951
DB 659 AEDKREKKGSRILKAKFLDK 678

RESULT 3
S57447
HPBR11-7 protein - human
N:Alternate names: HPBR11-4 protein
C:Species: Homo sapiens (man)
C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Oct-2004
C:Accession: S57447; S57489
R:Fleischhauer, K.L.
submitted to the EMBL Data Library, June 1992
A:Reference number: S57447
A:Accession: S57447
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-551 <FLE>
A:Cross-references: UNIPROT:Q16630; UNIPARC:UPI000006D566; EMBL:X67336; NID:g871300; PIDN:
A:Accession: S57489
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-551 <FL2>
A:Cross-references: UNIPARC:UPI000006D566; EMBL:X67337; NID:g871298; PIDN:CAA47752.1; PFI
C:Genetics:
A:Introns: 231/3
F:82-151/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match          4.2%; Score 316; DB 2; Length 551;
Best Local Similarity 28.0%; Pred. No. 2.1e-09;
Matches 111; Conservative 32; Mismatches 106; Indels 148; Gaps 21;

QY 19 PRGRGHGARGAPSPRPQNLRLHQQPPVQVQYEPSPAPSTTFNSPAPNF----LPP 74
DB 198 PQGGRGRGFPGA-----VPGGDRFGPAGP-----GGPPPPPPAGQTTP 237

QY 75 RPDVFVPPMPMPSAOGPLPCCPIRPP---FPHQMRHPPVY-----PPC--FPP 119
DB 238 RP---PLGPPGPPGPPPPGQVLPPLAGPNRGDRPPVLPFGQPGQPLGLPLP 294

QY 120 MPPMPCPNPPVPVPGAPGQ-----GTFP-----FMPPPPSMHPPPP--PPV 159
DB 295 GRPP--VPVGVGPPGPPPPGQGGPPPPPPPPPPGPPPLTLAPPHLPGPPGAPPP 353

QY 160 MPQVNYQPPGVSHNFPSPFNSFQNNPSPSLPSANSSSP---HFRHLPPY----- 210
DB 354 AP-HVN---PAF---FPPT-----NSGMTSDSRGPPPTDPYGRPPPYDRGDYG 396

QY 211 PLPKAFSERRSPERLKHVDHRRDHS----- 237
DB 397 PGREMDTARTPLUSEAFEEIWMNRRAISSAISRAVSADAGDYSATETLVTAISLIX 456

QY 238 -----HG-RGERHSLDRRERGRSPDRRRQDSRYRSYDRGRTP 275
DB 457 QSKVSADDRCKVLISSLQDCLHGIESKSGSGSRERSRERDHSR--SREKSRHKSRSR 514

QY 276 SRHSYRSRERERHRHRDNRRSPSLERSYKKEYK 312
DB 515 DRHDDYRERSRERHRDRD--RDRDRERDREREYR 549

RESULT 4
B69693
ribonuclease III (EC 3.1.26.3) - Bacillus subtilis
N:Alternate names: RNase D; RNase O
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69693; Jc4821
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
```


Db 176 IEYDILGETGPAHNKAFDAQVINGOVGLKSGRKTQKAEQSAQAQFAIN-----QLTHR 229

RESULT 7

T42727

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004

C:Accession: T42727

R:Witte, M.M.; Scott, R.E.

A:Title: DNA sequence of the EMBL Data Library, November 1998

A:Reference number: Z22246

A:Accession: T42727

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1560 <WIT>

A:Cross-references: UNIPROT:P97868; UNIPARC:UPI0000028767; EMBL:U83913; NID:g3858884; P1

A:Experimental source: strain Balb/C

C:Genetics:

A:Gene: P2P-R

C:Function:

A:Description: involved in hnRNP association and Rb1 binding

F:57-107/Domain: RING finger homology <RRN>

Query Match 3.9%; Score 293; DB 2; Length 1560;

Best Local Similarity 23.8%; Pred. No. 1.3e-07;

Matches 151; Conservative 72; Mismatches 200; Indels 212; Gaps 34;

QY 25 HGARPSA--PSRFPQNLRLHPQPPVQYQYEPSSA-----PSTTFNSPAPNPLP 73

Db 296 NAARPGGRPGWEHSN-KLGYLVSPQPIRRGERSCVSRINRGHSHSRSQRTQGPS-LP 353

QY 74 PRDFVPPPPMPPSAQGLPDCIPRPFPNQHMRPFPVPPCFPMPPMPPCNNPPVP 133

Db 354 ATPVFVFVPPPP-----PLYPPPP-----HTLPLP-----PGVLPPQSPQPP--P 391

QY 134 GAPPGGTPFPWMPPPSMHPHP-----PPVVMPPQVNV----- 166

Db 392 GQPPTAG---YSVPPGFPAPANISTACFSPGVPPTAHNTMTTQAPLLSREFFREQK 448

QY 167 -----QYP---PGYSHNPPPPPSFQNPSSFLPSANNSSPHRLPLPPYPLPKAPS 217

Db 449 RLKESKFPYVSGSSYSSTDSKSRSGTSRSYSRSGSRHSYSRSPVP----- 502

QY 218 ERSPERLKHYYDHRDHRDHSRGERHRSLSDRRCRSGSPDRRRQDSYSDYD-RGRTPS 276

Db 503 -RRRGKSRNY---RSRSRSHG---YHRS-----RSRSPYRRYHRSRSPQAQFGQSPT 550

QY 277 RHESYRSRERER-HRHR-----DNRSPSLERSYKKEYKR 313

Db 551 K-RNVPRG-ETERYFNRYREVPPVDIKAYGRSVDPRDPFEKERYERWERKYREWYEK 608

QY 314 SGRSYGLSVVPEPAGCTPELPGSIIKNNTDSWAP-----PLEIVNHRSPSREKKGARWEEK 369

Db 609 YYKGYAVGAQPRSA-----NREDFSPERLLPLNIRN--SPFTGREDYAAQ 655

QY 370 DRWSDN-----QSGGDKNYSIKEKPEETMPD-----KNE-----EE 403

Db 656 SHRNRLNGNYPEKLSRDSHNAKDNPKSKESESNVPGDGKGNKHKHGRKRENEEKGE 715

QY 404 EELLKPVMI-----RCHTSEN-----YSSDPMDQVGDSTVVGTSR 440

Db 716 SESFLNPELLGKPRKCRGSGSIDETKDTLTFVPPSRDATPVVRDEPMD--AESITPKSVS 773

QY 441 LRLDYKFEELGSRQEKAKA-----ARPPWPEPKTKLDELESSESE-----CES 487

Db 774 DKDKREKDKPKVSKDKTKRSGSATKDNVLPKSGKPGQKVDGDRKSPREPPPKKA 833

QY 488 DEPST-----CSSSSSDEVFDVIAIIRKKAH 514

Db 834 KEBATKIDSVKPSSSSQKDEKVTGT-----PRKAH 863

RESULT 8

B82073

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: B82073

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardon, D.; Emolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: B82073

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-235 <HEI>

A:Cross-references: UNIPROT:O9KPB2; UNIPARC:UPI0000134369; GB:AE004316; GB:AE003852; NID:

A:Experimental source: serogroup O1; strain Ni6961; biotype El Tor

C:Genetics:

A:Gene: VC2461

A:Map position: 1

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 285.5; DB 2; Length 225;

Best Local Similarity 32.2%; Pred. No. 2.8e-08;

Matches 75; Conservative 42; Mismatches 101; Indels 15; Gaps 5;

QY 1101 SPVLOKLTFFEAIGVIFTHVLLARAFTLRTVGVNHLTLGHNORMEFLGDSIMQLVATE 1160

Db 2 TTPMKNLT---SKLGYTFKETELNLALTHRSANGK-----HNERLEFLGDSILSFVIAD 53

QY 1161 YLFIHPKDDHGHGTLRLSSLVNNTQAKVAELGMOEYAITNDKTRPVGLRTKT-LAD 1219

Db 54 EYRRPFPKNGEDMRATLVGNVTLAEGLREFDLGDLVILGPGELKSGGFRDSILAD 113

QY 1220 LLESFAALYTDKLELYVHTFMNVCFPPRLKBFILNQDWNPKSQIQCCQLTLRTGKBP 1279

Db 114 AVEAIIAGIYDSDLETARSIVLEWYHGLEIKPGASQKDPKRLQEF-----LQGRKK 168

QY 1280 DPLPYKLTQVGPSHARTVTVAVPKG-ERICGGKGPSIQQAEMGAMDALEK 1331

Db 169 PLFVYTVTNIKGEAHNQEFTVACEVAGMDTPVIGKTSRRKAEQAAATALEQ 221

RESULT 9

A83961

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: A83961

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: A83961

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-263 <STO>

A:Cross-references: UNIPROT:O9KA05; UNIPARC:UPI0000134340; GB:AP001515; GB:BA000004; NID:

A:Experimental source: strain C-125

C:Genetics:

A:Gene: rncS

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 282; DB 2; Length 263;

Best Local Similarity 30.9%; Pred. No. 5.2e-08;

Matches 77; Conservative 50; Mismatches 110; Indels 12; Gaps 5;

QY 1084 HPLQLQEPNTDQLIETSPVLOKLTFFEAIGVIFTHVLLARAFTLRTVGVNHL--LTLG 1141

Db 15 HSERRFQ---KRLTLTAKQQQMFDELLTLNLTLPANKLLVQAFTHSSVTNEHRTQSCK 71

QY 1142 HNORMEFLGDSIMQLVATEYLFTHFPDHHGHGTLRLSSLVNNTQAKVAELGMOEYAI 1201

Db 532 GSTRSRYSRFSRSHRSYSRSPYP-----RRGRGKSRNY-----RSKRSHG---YH 578
QY 245 RSLDRRGRSPPRRQDRGRYRDYD-RGRTPSRHSYRSRER----- 289
Db 579 RS-----RSRSPRYRHSRSRSPQAFRGQSPNKNVPOGETERYFNRYREVPYPMK 633
QY 290 -----ERRHRDNRRSPLERSYKKEYKRSYGLSVVPPAGCTPELPG 336
Db 634 AYYGRSVDFRDPPEKERYE-----WEKYRWYKYYKGYAAGQPRSA----- 679
QY 337 IIKNTDWSAP-----PLEIVNHRSPSRKKRARWEEKDRWSDNQSSG-----KD 381
Db 680 ---NRENFSRERPLNLNRN--SPFTRGRREDVVGQSHRSRNIAGNYPEKLSARDGHQ 734
QY 382 KNYTSIKEKEPESTMPD-----KNBEEBELKPVWIRCTHSENYSSDPM-- 427
Db 735 KONTSKEKESENAPGDGKGNKHKKRKGEGSEGLNPELLETSRK-----SREPTGV 790
QY 428 -DQVGDSTVVGTSR-----LRDLYDKFEE---LGSRQEKAKAARPPWPPKTKLDE 478
Db 791 EENKTDLSLFLPRDDATPVRD--EPMDAESITFKSVSEKDKRER---DKPKAKGDKTKR 845
QY 479 SSESCECEDSTCSSSDSE-----VFDVIAEIKRKAHPDRL 518
Db 846 KNDGSAVSKKENIVKPAKGQKVDGVRDLDLNLQLKPKERLRL 893
RESULT 13
T06291
extensin homolog T9E8.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06291
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuel
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15588
A:Accession: T06291
A:Molecule type: DNA
A:Residues: 1-760 <BEV>
A:Cross-references: UNIPROT:Q3TOK5; UNIPARC:UPI000009FB81; EMBL:AL049608
A:Experimental source: cultivar Columbia; BAC clone T9E8
C:Genetics:
A:Map position: 4
A:Note: T9E8.80
Query Match 3.7%; Score 275; DB 2; Length 760;
Best Local Similarity 35.1%; Pred. No. 4.7e-07;
Matches 88; Conservative 15; Mismatches 96; Indels 52; Gaps 12;
QY 14 PGRGCPRGCGHGRSAPSFRQNL-----RLHPOQPPVQYVEPPSAPS----- 60
Db 361 PGRPAQRSPGCKAFLSRP---PVCNCSFSCGRSVSP-RPPVVTPLPPLPSLPSPPPAP 416
QY 61 ----TTFNSNAPNFPRLPRDF-VPPPPMPPSAQGLPRLPCPTRRPPFNHGMHRPPVP 115
Db 417 FSTPPTLTSPDPS--PPPPVVSPPPPPPPPVVSPPPPPPPPPPVVSPPPPPPPP 474
QY 116 CFPFM---PPMPCPNPPVFGAPPQGTFPFMPPPSMHPHPPPPV-----MPQ 162
Db 475 PPPVVSPPPPPPPPPPVVSPP--PPPPPPPPVVSPPPPPPVVSPPPPPPSPAPT 530
QY 163 QVNYQVPPGVSHNFPPPSPNSFQNNP--SSFLPSANNSSPH-----RHLPVPL----- 212
Db 531 PVYCTRPPPPHSPPPPPQSPPPPPVYSSPPPPHSPPPHSPPPHSPPPPPPIYPLS 590
QY 213 -PKAPERRSP 222
Db 591 PPPPPTPVSSP 601
RESULT 14
T48946

hypothetical protein T15B3.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T48946
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.; I
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25009
A:Accession: T48946
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1531 <JOR>
A:Cross-references: UNIPROT:Q9LXW7; UNIPARC:UPI000009DAF8; EMBL:AL1163975; GSPDB:GN000061;
A:Experimental source: cultivar Columbia; BAC clone T15B3
C:Genetics:
A:Gene: ATSP.T15B3.60
A:Map position: 3
A:Introns: 45/2; 100/3; 138/3; 183/3; 227/3; 278/3; 322/1; 357/2; 421/3; 475/2; 584/1; 6
Query Match 3.7%; Score 274.5; DB 2; Length 1531;
Best Local Similarity 21.3%; Pred. No. 1.2e-06;
Matches 247; Conservative 142; Mismatches 420; Indels 351; Gaps 58;
QY 372 WSDNQSGKDKNYTSIKEKEPESTMPDKNEBEEELKPVWIRCTHSENYSSDPMDQVG 431
Db 455 YSQSQKHAKQSNKXSIMFL--ERGNPKQRDHLHLMRREVLIQDPEAPNLKSCPPVYKNG 512
QY 432 DST-VVGTSLRLDYDKFEEELGSRQEKAKAARPPWE---PP--KTKLDED-LESSSES 483
Db 513 HGVKEIGSMVVPDSNITVSEEAATQMSD---PPSRNEQLPCKKLRDNNLLQSNKGE 569
QY 484 ECESDESDTSCSSSDSEVFDVIAEIKRKAHPDRLHDEL---WYNDPGQMNQDGP----- 534
Db 570 KVASSKSKSSSAAGSK-----KKELHGTTCANALSGTW---GENIDGATQAYK 617
QY 535 ---LCKCSAK-----ARRTG-IRHSIYPERAIAK---PCRP----- 564
Db 618 FDFCCNISGEVYSSFLLESTLAEDVGKVEDMLVLRKLVKASVSPCGQIRLSQEELVK 677
QY 565 -----TNNAGRLFYRITVSPPTNFLTDRPTVIEYDDHEYIEFGSMFAHPLTNI 615
Db 678 AKYFOOFFNMGFKLVGSKSQGTKEFLLQ-----TDTSSLWHDPAMFLLLPVETN 730
QY 616 PLCKVIRFNIDYTHIEEMMPENFCVGLGELFSLFRD-ILELYDMNLKGLPFDSP 674
Db 731 DLAS--SATDWS-----AINSCASIVE---FLKKNSLDLRD-----SDGNQ 768
QY 675 CCFRHFMRFRVRFLPDGQKVLSMHQILLYLLRCSKALVPBEEIANNLOWEELEWQKVA 734
Db 769 C-----NTSSQEVL-----LDDKMEETNLHFANASSDKNS 800
QY 735 BECKGMIVNPGTKPSSVRI--DQDREOFNPDVITFPIIVHFGIRPAQLSYAGDPQYOK 792
Db 801 LEELVIAIHTGRISYIWEAVSDSSAMSPFEVD-----ASSGYATYAEYFN 846
QY 793 LWKSY-VKLRHLIANSF--KVQTDKQK---LAQREELQKIRQK--NTWRREVTVELSS 844
Db 847 --KKYGIVLAH--PNQPLMKLQSHHAHNLVDVFNEMVKTEPKAGNVRKRPNI--- 898
QY 845 QGFWKTGIRSDVCQHAMLFPVTHIRYHQCLMHLDKLIGYTPQDRCLQLAMTHFS--H 902
Db 899 -----HAHLPEL-----LARID-----VPRAVLKSIIYLLSPVMH 928
QY 903 HLNFGMNPDPHARNLSNCGIRPKYCDRKVHHMHMKKGINTLINIMSLRGQDDPTPSRI 962
Db 929 RLESMLASQLREEI-DCSIDNFSISSTSI-----LEAVTTL-----TCPES 969
QY 963 NHNERLEFLGDAAVVEFLTSVHLVYLPFSLSEEGGLATYRTAIYVONQHAML--AKKLE--- 1017
Db 970 FSWERLELLGDSVLKVASCHLEPLKYPDKDEGQLSRQSIISNHLRLTTSEKLQGYI 1029
QY 1018 ----LDPFMLYAHG-----PDLG-----RESDLRH----- 1038
Db 1030 RNGAFERPRWTAPGQPSLFPVPCCKGIDTRVPLDPKPFETENMTIKIGKSCDMGRWVVS 1089

Qy	1039	AMANC	EALIGAV	YLGSGS	EEAKQL	FGRLFN	---	DPD	LR	EVW	LN	YPL	HLPL	QLQ	QEN	TND	1094																																						
		:::		:	:	:	:	:	:	:	:	:	:	:	:	:																																							
Db	1090	KSVSD	CAEALIG	AYVSGG	SASLHM	KMKL	GID	VD	PNL	VVE	AIN	RV	---	SL	RCY	IPKED	1147																																						
Qy	1095	ROLI	ETSPV	QKLTE	PEA	IGVITH	VR	LLA	RA	FT	LV	GEN	HL	TGHN	---	QRM	FLGD 1151																																						
Db	1148	-----	EL	TEL	ERKI	QHEF	S	KFLP	L	KEA	T	-----	HSS	RES	YSY	ERLE	FLGD 1189																																						
Qy	1152	SIMOL	VATE	YLF	FI	HPD	H	HEG	HL	TL	LR	SSL	V	NN	TQ	AKVAE	-----ELGM 1196																																						
Db	1190	SVLDF	L	TR	HL	FN	TY	ET	Q	PE	M	T	LR	S	C	VNN	ENFAQVAKNNLHTLORCATVLETOI 1249																																						
Qy	1197	Q	EY	AIT	---	NDK	TKR	---	PV	G	L	R	T	K	L	AD	LESFIAALYTKOLEYVHTFMNVCFPPRLKE 1251																																						
Db	1250	NDYLS	FQK	PE	DE	CR	S	I	P	S	I	Q	G	P	A	L	GDUVESIAGALLITRLDLDQVWR--VPEPLLSP 1307																																						
Qy	1252	FILNQ	DWN	-	DPK	SOL	QO	CC	U	T	L	---	R	T	E	G	K	E	P	D	I	P	L	Y	T	K	L	Q	-----TWG	PSHARTYT 1299																									
Db	1308	L	V	P	D	K	L	Q	L	P	P	Y	R	E	L	N	E	L	C	D	S	L	G	Y	F	F	R	V	K	C	S	N	D	G	V	K	A	Q	I	Q	L	D	D	V	L	T	G	D	G	S	E	Q	T	---	1365
Qy	1300	V	A	Y	F	K	E	R	I	G	C	G	K	P	S	I	Q	A	E	M	A	D	A	L	E	K	N	F	P	O	M	A	H	K	R	P	I	G	R	Y	-----RQEL	KEM 1355													
Db	1366	-----	NK	L	A	G	K	A	A	S	H	L	L	T	Q	---	E	K	R	N	I	S	R	K	T	S	L	G	D	N	Q	S	M	D	V	1401																			
Qy	1356	R	W	E	R	H	O	E	R	---	P	D	E	T	E	D	I	K	1373																																				
Db	1402	N	L	A	C	N	H	S	D	R	E	T	L	T	S	E	T	T	E	I	Q	1421																																	

RESULT 15

G86292
hypothetical protein F7H2.17 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: G86292
R/Theologis, A.; Ecker, M.K.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marshall,
Rizzo, M.H.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: G86292
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1006 <STO>
A/Cross-references: UNIPROT:Q9LMQ1; UNIPARC:UPI0000002407; GB:AE005172; NID:g8927662; P
C/Genetics:
A/Map position: 1

Query Match 3.6%; Score 273.5; DB 2; Length 1006;
Best Local Similarity 32.9%; Pred. No. 8e-07;
Matches 99: Conservative 17; Mismatches 98; Indels 87; Gaps 19;

Qy	3	QGNTCHRMSTFHGRCGRG-----GHGARPSAPSFQNRLRLHHPQP	47
Dd	48	QSNRGRESS----GSSNRGRSCDPLQYLFDTCGHWPPPTTS--PENPFL--PRQP	98
Qy	48	PVQYQYEPPSPASTTFSNSAPNFLPRRDFVP-FPPMPPSAQGLPCPIRRPFPHNQ	106
Dd	99	P-----RPPRP-----PRPRSRLPPLVSPSPPLHRP--SFCPPPL	138
Qy	107	KRHFPFVPPCPFPMPMPCPNNPVPGAPCQGTFP-----FMMPSPSMHPHPPPPVM	160
Dd	139	MPSPEPLVPS-PPPPPSELVSP-PSPPPPFFPSPBPVIYFPELVPSPPPELPG	196
Qy	161	PQQVNYQVPGCYSHNFPPPPFSGFNQNPSSFLPSANNSSGFHFHLPP---YPLPKAPS	217

197	GDQTT-QEPPLW----	LPPLPPGD-ETPVSLPPPLDEFPF----	MPPIWLPPDPVA	246
Db				
218	ERRSPERLKH-----	YDHRHDHSHGCRGHRHSLSRRRGRGSPDRRRQDS	263	
Qy				
247	QTSAEAFDQPLPLVTITEAPNPHNSHRHD-----	ENKKGLODRNR-RVKSRRRRRS	299	
Db				
264	R 264			
Qy				
300	R 300			
Db				

Search completed: December 24, 2005, 00:54:08
Job time : 33 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 11:30:58 ; Search time 2599 Seconds

(without alignments)
12216.466 Million cell updates/sec

Title: US-10-774-974-1

Perfect score: 4764

Sequence: 1 cgtctgtacctgcggtgta.....aaaaaaaaaaaaaaaaaaaaa 4764

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4764	100.0	4764	9	ACD27526
2	4764	100.0	4764	12	ADQ96674 Human rib
3	4764	100.0	4764	14	AEb47410 Human rib
4	4760.8	99.9	4764	10	ADG32900 Human DNA
5	4760.8	99.9	4764	12	ADQ17463 Human sof
6	4760.8	99.9	4764	13	ADf68726 Human CDN
7	4760.8	99.9	4764	14	ADX03768 Human CDN
8	4760.8	99.9	4764	14	ADY17385 DNA encod
9	4732.4	99.3	5425	12	ADQ22183 Human sof
10	3376.2	70.9	3603	14	AEa19628 Novel hum
11	2582	54.2	2598	4	AaH13924 Human CDN
12	1735.6	36.4	1784	14	AEa20644 Novel hum
13	1619.6	34.0	1626	9	ACf25385 Human put
14	1439.8	30.2	1455	4	AaS33094 DNA encod
15	1398.6	29.4	1458	4	AaS34779 CDNA enco
16	1398.6	29.4	1458	10	ADc45937 Human neo
17	1262.4	26.5	1314	4	AaS33296 DNA encod
18	1262.4	26.5	1314	4	AaS34943 CDNA enco
19	1262.4	26.5	1314	10	ADc46101 Human neo

20	818.8	17.2	897	4	AAH04336	Aah04336 Human cdn	
21	770.2	16.2	4275	4	ABL03445	AbI03445 Drosophil	
22	725.4	15.2	1003	4	AAF22509	Aaf22509 Human bre	
23	724	15.2	803	4	AAF22579	Aaf22579 Human bre	
c	24	703.8	14.8	6397	4	ABL03444	AbI03444 Drosophil
	25	683.4	14.3	833	4	AAF22581	Aaf22581 Human bre
c	26	602.4	12.6	763	4	AAF22510	Aaf22510 Human bre
c	27	581.6	12.2	701	4	AAF22580	Aaf22580 Human bre
28	579.8	12.2	591	14	ACL59694	ACl59694 Human col	
29	569.6	12.0	744	10	ADD34804	ADd34804 Mouse mit	
30	518.2	10.9	578	2	AAZ17689	Aaz17689 Human gen	
31	469.4	9.9	633	5	ABV26800	Abv26800 Human pro	
32	469.4	9.9	633	5	ABV20994	Abv20994 Human pro	
33	469.4	9.9	633	5	ABV20950	Abv20950 Human pro	
34	469.4	9.9	633	5	ABV26840	Abv26840 Human pro	
c	35	458.2	9.6	501	4	AAH09845	Aah09845 Human cdn
	36	442.4	9.3	5828	5	ABA19864	AbA19864 Human ner
37	442.4	9.3	32223	5	ABA17368	AbA17368 Human ner	
38	442.4	9.3	32223	5	ABA19863	AbA19863 Human ner	
39	413.4	8.7	9928	5	ABA19865	AbA19865 Human ner	
40	411.4	8.6	488	9	ACH33094	ACH33094 Human end	
41	405.4	8.5	583	5	ADL45324	ADl45324 Human ova	
42	399.2	8.4	3809	6	ABL53489	AbI53489 Human non	
43	392.2	8.2	405	4	AA537351	Aa537351 Novel hum	
44	384	8.1	397	3	AAH30538	Aah30538 Human col	
45	381.8	8.0	394	5	AAF66326	Aaf66326 Novel hum	

ALIGNMENTS

RESULT 1
ACD27526
ID ACD27526 standard; cdna; 4764 BP.

XX ACD27526;

DT 18-SEP-2003 (first entry)

DE Human cdna encoding RNase III.

XX Human; ss; gene; RNase III; double stranded RNase; RNA target;
XX Gene silencing.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 246..4370

FT /*tag= a

FT /product= "RNase III"

FT /transl_except= (pos:297..299,aa:Cys)

FT /transl_except= (pos:3876..3878,aa:Gly)

FT /transl_except= (pos:4278..4280,aa:Gly)

XX US2003044941-A1.

XX 06-MAR-2003.

XX 20-FEB-2002; 2002US-00079185.

XX 06-JUN-1996; 96US-00659440.

XX 06-JUN-1997; 97US-00870608.

XX 07-JAN-2000; 2000US-00479783.

XX 06-JUL-2001; 2001US-00900425.

XX (CROO/) CROOKE S T.

XX Crooke ST;

XX WPI: 2003-521756/49.

XX P-PSDB; ABU63361.

Eliciting a modification of a selected RNA target in a cell, useful for

promoting inhibition of gene expression in a cell, comprises contacting an RNA-like polynucleotide-RNA target duplex with a polypeptide having an RNase III domain.

Example 2: Page 13-15: 33pp; English.

The invention relates to eliciting a modification of a selected RNA target in a cell comprises contacting an RNA-like polynucleotide-RNA target duplex with a polypeptide comprising an RNase III domain. Also included are promoting gene silencing in a cell, inhibiting the expression of a gene in a cell comprising employing the method of cited above, promoting inhibition of expression of a gene, a hybrid RNase III (comprising at least one domain from a human RNase III and at least one domain from an RNase III of an organism other than human) and a cell having enhanced RNase III activity over an activity exhibited by a second cell (where the second cell is not enriched with respect to the amount or activity of RNase III polypeptide). The method is useful for eliciting a modification of a selected RNA target in a cell, and for promoting inhibition of expression of a gene in a cell. Compositions comprising RNase III polypeptides or polynucleotides are useful for research, biological and clinical purposes. The polynucleotides may be used in defining the roles of RNase III and the interaction of human RNase III and cellular RNA. Host cells can be used for the production of human RNase III and for identifying agents, which increase or decrease levels of expression or activity of human RNase III in the cell. The present invention encodes human RNase III (a double stranded RNase).

Db ||||| 1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACTGTTCTAGAGCTCAGACT 1740
Qy ||||| 1741 CTGAAAGTTTTGACGTTATTGACAGAAATCAAAACGCAAAAAGGCCACCTGACCGACTTC 1800
Db ||||| 1741 CTGAAAGTTTTGACGTTATTGACAGAAATCAAAACGCAAAAAGGCCACCTGACCGACTTC 1800
Qy ||||| 1801 ATGATGAACCTTTGGTACACGATCCAGGCCAGATGAATGATGACCACTCTGCAATGCA 1860
Db ||||| 1801 ATGATGAACCTTTGGTACACGATCCAGGCCAGATGAATGATGACCACTCTGCAATGCA 1860
Qy ||||| 1861 GGGCAAAAGCAAGACGACAGCAAGAAATAGGACAGCAATTTATCTCTGGAGAAAGGCCATCA 1920
Db ||||| 1861 GGGCAAAAGCAAGACGACAGCAAGAAATAGGACAGCAATTTATCTGAGAAAGGCCATCA 1920
Qy ||||| 1921 AGCCCTGTCTCTATGACCAACAATGCTGGCAGACTTTTCCACTACCGGATCAGACTCT 1980
Db ||||| 1921 AGCCCTGTCTCTATGACCAACAATGCTGGCAGACTTTTCCACTACCGGATCAGACTCT 1980
Qy ||||| 1981 CCCGCCCTACGAACTTTTAACTGACAGCCCAACTGTTATAGNAATACGATGATCAGAGT 2040
Db ||||| 1981 CCCGCCCTACGAACTTTTAACTGACAGCCCAACTGTTATAGNAATACGATGATCAGAGT 2040
Qy ||||| 2041 ATATCTTTCAAGGATTTCTATGTTTGACATGCCGCCCTGACCAATATTTCCACTGTGTA 2100
Db ||||| 2041 ATATCTTTCAAGGATTTCTATGTTTGACATGCCGCCCTGACCAATATTTCCACTGTGTA 2100
Qy ||||| 2101 AAGTAATAGATTCAACATAGACTACACGATTCATTTCAATGAAGAGATGATGCCGGAGA 2160
Db ||||| 2101 AAGTAATAGATTCAACATAGACTACACGATTCATTTCAATGAAGAGATGATGCCGGAGA 2160
Qy ||||| 2161 ATTTTGTGTGAAGGGCTTGAACCTTTTCACTGTGTTCTTATTCAGAGATATTTTGGAAAT 2220
Db ||||| 2161 ATTTTGTGTGAAGGGCTTGAACCTTTTCACTGTGTTCTTATTCAGAGATATTTTGGAAAT 2220
Qy ||||| 2221 TATATGACTGGAACTTTAAAGTTCCTTTGTAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
Db ||||| 2221 TATATGACTGGAACTTTAAAGTTCCTTTGTAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
Qy ||||| 2281 TTCAATTTCAATGCGCAATTTTGAAGATTTCTTCCAGATGAGGAAAGGAAAGTCTGTCCA 2340
Db ||||| 2281 TTCAATTTCAATGCGCAATTTTGAAGATTTCTTCCAGATGAGGAAAGGAAAGTCTGTCCA 2340
Qy ||||| 2341 TGCACAGATTTCTCTGTAATTTGTAAGGTGAGCAAAAGCCCTGTCGTCGAGGAGGAGA 2400
Db ||||| 2341 TGCACAGATTTCTCTGTAATTTGTAAGGTGAGCAAAAGCCCTGTCGTCGAGGAGGAGA 2400
Qy ||||| 2401 TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGACAGAAATGCAAG 2460
Db ||||| 2401 TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGACAGAAATGCAAG 2460
Qy ||||| 2461 GCATGATTTGTTACCAACCTTGGGAGGAGGAGCTGGAGTGGCAGAAATATGACAGAAATGCAAG 2520
Db ||||| 2461 GCATGATTTGTTACCAACCTTGGGAGGAGGAGCTGGAGTGGCAGAAATATGACAGAAATGCAAG 2520
Qy ||||| 2521 GTGAACAGTTCAACCCCGATGTAATTTTCCGATTTATGTCATTCGATTCGATCACTGGATC 2580
Db ||||| 2521 GTGAACAGTTCAACCCCGATGTAATTTTCCGATTTATGTCATTCGATTCGATCACTGGATC 2580
Qy ||||| 2581 CTGCAAGTTGATTTGACAGAGACCCACAGTACCAAAATCTGTGGAGAGCTTATGTA 2640
Db ||||| 2581 CTGCAAGTTGATTTGACAGAGACCCACAGTACCAAAATCTGTGGAGAGCTTATGTA 2640
Qy ||||| 2641 AACTTCGCCACCTCTTAGCAAAATAGTCCCAAAAGTCAAACTGACAAACAGAAAGCTGG 2700
Db ||||| 2641 AACTTCGCCACCTCTTAGCAAAATAGTCCCAAAAGTCAAACTGACAAACAGAAAGCTGG 2700
Qy ||||| 2701 CACAGAGGAGGAGGAGCCCTTCAAAAATACGGCAGAGAAATCAATGACAGCAAGTAA 2760
Db ||||| 2701 CACAGAGGAGGAGGAGGAGCCCTTCAAAAATACGGCAGAGAAATCAATGACAGCAAGTAA 2760
Qy ||||| 2761 CGGTGAGCTAATAGTACCAAGATTTCTGAAATCTGGCAATCCGTTCTGTATGCTGTACAGC 2820

Db ||||| 2761 CGGTGAGCTAATAGTACCAAGATTTCTGAAATCTGGCAATCCGTTCTGTATGCTGTACAGC 2820
Qy ||||| 2821 ATGCAATGATGCTACTGTTCTGACCCATCATATCCGCTACCAACCAATGCTTAATGCAAT 2880
Db ||||| 2821 ATGCAATGATGCTACTGTTCTGACCCATCATATCCGCTACCAACCAATGCTTAATGCAAT 2880
Qy ||||| 2881 TGGACAAGTTGATAGGATATCTTTTCCAAGATGCTTTGTCAGCTGGCCATGACTC 2940
Db ||||| 2881 TGGACAAGTTGATAGGATATCTTTTCCAAGATGCTTTGTCAGCTGGCCATGACTC 2940
Qy ||||| 2941 ATCCAAGTCACTATTTAAATTTTGGAAATGAAATCTCTGATCATGCCAGAAATTTCAATCTA 3000
Db ||||| 2941 ATCCAAGTCACTATTTAAATTTTGGAAATGAAATCTCTGATCATGCCAGAAATTTCAATCTA 3000
Qy ||||| 3001 ACTGTGGAAATTCGGCAGCCCAAAATACGGAGACAGAAAAGTTTATCATCATGCAATCGGA 3060
Db ||||| 3001 ACTGTGGAAATTCGGCAGCCCAAAATACGGAGACAGAAAAGTTTATCATCATGCAATCGGA 3060
Qy ||||| 3061 AGAAAGGGATTAACACCTTTGATAAATATCATGTACAGCTTGGCCAAAGATGACCCAACTC 3120
Db ||||| 3061 AGAAAGGGATTAACACCTTTGATAAATATCATGTACAGCTTGGCCAAAGATGACCCAACTC 3120
Qy ||||| 3121 CCTCGAGGATTAACCAATGAAACGTTTGGAAATTCCTGGGTGATGCTGTTGTTGAATTC 3180
Db ||||| 3121 CCTCGAGGATTAACCAATGAAACGTTTGGAAATTCCTGGGTGATGCTGTTGTTGAATTC 3180
Qy ||||| 3181 TGAACAGCTCCTATTTGATCTATTTGTTTCTTAGTCTGGAAGAGAGGATTTAGCAACCT 3240
Db ||||| 3181 TGAACAGCTCCTATTTGATCTATTTGTTTCTTAGTCTGGAAGAGAGGATTTAGCAACCT 3240
Qy ||||| 3241 ATCGGACTGCCATTTTGAAGCTTGAACCTTTGAGAGAAATCGGACCTTCGACATG 3300
Db ||||| 3241 ATCGGACTGCCATTTTGAAGCTTGAACCTTTGAGAGAAATCGGACCTTCGACATG 3300
Qy ||||| 3301 ATCCATTTATGCTGATGCTCAACGCGCTTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
Db ||||| 3301 ATCCATTTATGCTGATGCTCAACGCGCTTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
Qy ||||| 3361 CAATGCCAAATTTTGAAGCTTAAATAGGAGCTGTTTACTTGGAGGAAAGCTCGGAG 3420
Db ||||| 3361 CAATGCCAAATTTTGAAGCTTAAATAGGAGCTGTTTACTTGGAGGAAAGCTCGGAG 3420
Qy ||||| 3421 AGCCAAAGGATTTTGGACGCTTCTTTAAATGATCCGACCTGCGGAGTCTGCG 3480
Db ||||| 3421 AGCCAAAGGATTTTGGACGCTTCTTTAAATGATCCGACCTGCGGAGTCTGCG 3480
Qy ||||| 3481 TCAATTTATCTCTCCACCCACTCCAACTCAAGAGCCAAATCTGATCGACAACTTTATTG 3540
Db ||||| 3481 TCAATTTATCTCTCCACCCACTCCAACTCAAGAGCCAAATCTGATCGACAACTTTATTG 3540
Qy ||||| 3541 AAATCTCTCCAGTTCTACAAAACTTACTGAGTTTGAAGAACTTGGAGTAAATTTTAA 3600
Db ||||| 3541 AAATCTCTCCAGTTCTACAAAACTTACTGAGTTTGAAGAACTTGGAGTAAATTTTAA 3600
Qy ||||| 3601 CTCATGTTGACTTTTGGCAAGGCAATTCATGAGTCTGGGATTTTAACTCTGA 3660
Db ||||| 3601 CTCATGTTGACTTTTGGCAAGGCAATTCATGAGTCTGGGATTTTAACTCTGA 3660
Qy ||||| 3661 CCTAGGCGCAATCAGAGAAATGGAATTCCTAGGTGACTTCCATATGCAACTTGGTAGCCA 3720
Db ||||| 3661 CCTAGGCGCAATCAGAGAAATGGAATTCCTAGGTGACTTCCATATGCAACTTGGTAGCCA 3720
Qy ||||| 3721 CAGAGTACTTATTTCAATTTTCCAGATCATCATGAAGGACACTTAACTTTGTTGCGAA 3780
Db ||||| 3721 CAGAGTACTTATTTCAATTTTCCAGATCATCATGAAGGACACTTAACTTTGTTGCGAA 3780
Qy ||||| 3781 GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGGTTAGCGGAGGAGTGGGCAATGCAAGT 3840
Db ||||| 3781 GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGGTTAGCGGAGGAGTGGGCAATGCAAGT 3840
Qy ||||| 3841 ACGCCATTAACCAACGACCAAGAGCCCTGTCGCTTCCGACCAAGACCTTGGCGG 3900
Db ||||| 3841 ACGCCATTAACCAACGACCAAGAGCCCTGTCGCTTCCGACCAAGACCTTGGCGG 3900

QY 3901 ACCTTTTGGAAATCATTTATTGACGCGCTGTACACTGATAAGGATTTGGAAATATGTTCTATA 3960
DB |||||||
QY 3901 ACCTTTTGGAAATCATTTATTGACGCGCTGTACACTGATAAGGATTTGGAAATATGTTCTATA 3960
DB |||||||
QY 3961 CTTTTCATGAATGTCGCTCTCTTTTCCACGAGTTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
DB |||||||
QY 3961 CTTTTCATGAATGTCGCTCTCTTTTCCACGAGTTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
DB |||||||
QY 4021 ATGACCCCAAAATCCAGCTTTACGAGTGTGCTTGACACTTAGGACAGAAAGAAAGAGC 4080
DB |||||||
QY 4021 ATGACCCCAAAATCCAGCTTTACGAGTGTGCTTGACACTTAGGACAGAAAGAAAGAGC 4080
DB |||||||
QY 4081 CAGACATTCCTCTGTACAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTCA 4140
DB |||||||
QY 4081 CAGACATTCCTCTGTACAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTCA 4140
DB |||||||
QY 4141 CTGTGGCTGTTTATTTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGACCAAGTATTTCAGC 4200
DB |||||||
QY 4141 CTGTGGCTGTTTATTTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGACCAAGTATTTCAGC 4200
DB |||||||
QY 4201 AAGCGGAAATGGAGCAGCAATGGATGCGCTTGAAAAATATATTTTCCCGAGATGGCCC 4260
DB |||||||
QY 4201 AAGCGGAAATGGAGCAGCAATGGATGCGCTTGAAAAATATATTTTCCCGAGATGGCCC 4260
DB |||||||
QY 4261 ATCAGAGCGGTTTCATCGAACGAGTACAGACAGAGTTTAAAGAAATGAGTGGGAA 4320
DB |||||||
QY 4261 ATCAGAGCGGTTTCATCGAACGAGTACAGACAGAGTTTAAAGAAATGAGTGGGAA 4320
DB |||||||
QY 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATTAAGAGAGGCGAT 4380
DB |||||||
QY 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATTAAGAGAGGCGAT 4380
DB |||||||
QY 4381 GCAAGTGTGGAGTATTACTGTGCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
DB |||||||
QY 4381 GCAAGTGTGGAGTATTACTGTGCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
DB |||||||
QY 4441 AGTTTTCCTGCAGACATCAAGCAAGTGTGCTCAATGAAATTAAGAAATACAGAGTCAATCG 4500
DB |||||||
QY 4441 AGTTTTCCTGCAGACATCAAGCAAGTGTGCTCAATGAAATTAAGAAATACAGAGTCAATCG 4500
DB |||||||
QY 4501 CTATTGTTGTTTAAATGATCTGTTTATTAGCTGATGCTCTTTATTACAAAGTATTAGATT 4560
DB |||||||
QY 4501 CTATTGTTGTTTAAATGATCTGTTTATTAGCTGATGCTCTTTATTACAAAGTATTAGATT 4560
DB |||||||
QY 4561 TTTCTTCTATTAAACGGAAATCTGACTTTGGTGAATGTCATCTCTTTTATTG 4620
DB |||||||
QY 4561 TTTCTTCTATTAAACGGAAATCTGACTTTGGTGAATGTCATCTCTTTTATTG 4620
DB |||||||
QY 4621 CTCTTTAAATATAAATTCAGAGCATATTTCTATGTGGATAGATCTGTTTTCAT 4680
DB |||||||
QY 4621 CTCTTTAAATATAAATTCAGAGCATATTTCTATGTGGATAGATCTGTTTTCAT 4680
DB |||||||
QY 4681 CTGTGCCCCAGATTTGACCTAGACTTTTCAATTTGACAAATGAAAAATTTGACTTTACTAG 4740
DB |||||||
QY 4681 CTGTGCCCCAGATTTGACCTAGACTTTTCAATTTGACAAATGAAAAATTTGACTTTACTAG 4740
DB |||||||
QY 4741 TAAAAAATTT 4764
DB |||||||
QY 4741 TAAAAAATTT 4764
DB |||||||

RESULT 2

ADQ96674
ID ADQ96674 standard; cdna; 4764 BP.

XX AC ADQ96674;

XX XX 23-SEP-2004 (first entry)

DE Human ribonuclease III (RNase III) cdna.

XX Human; ribonuclease III; RNase III; research purpose; biological purpose;

clinical purpose; gene; cellular interaction; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 246..4370
/*tag= a
/product= "Human ribonuclease III protein"
/transl_except= (pos:297..299, aa:Cys)
/transl_except= (pos:3876..3878, aa:Gly)
/transl_except= (pos:4278..4280, aa:Gly)

PN US2004126867-A1.

PD 01-JUL-2004.

PF 09-FEB-2004; 2004US-00774974.

PR 06-JUL-2001; 2001US-00900425.

XX (CROO/) CROOKE S T.
XX (WUHH/) WU H.

PI Crooke ST, Wu H;

XX WPI: 2004-516913/49.

DR P-PSDB; ADQ96675.

DR GENBANK; AF189011.

XX New isolated nucleic acid molecule encoding human RNase III, useful for

PT research, biological, or clinical purposes, e.g. defining the roles of

PT RNase III and the interaction of human RNase III and cellular RNA.

XX Claim 2; SEQ ID NO 1; 31pp; English.

XX The invention relates to human ribonuclease III (RNase III) and its
CC corresponding nucleic acid sequence. The polynucleotide sequence of the
CC invention is useful for research, biological and clinical purposes. It is
CC useful in defining the roles of RNase III and the interaction of human
CC RNase III and cellular RNA. The present sequence is human RNase III cDNA.

XX Query Match 100.0%; Score 4764; DB 12; Length 4764;

XX Best Local Similarity 100.0%; Pred No. 0;

XX Matches 4764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCTGTGTACCTGCGGTAGTAGCTGCTTGTCTCTGACGCGCATCTCGGCGCCGAG 60

DB 1 CTGCTCTGTGTACCTGCGGTAGTAGCTGCTTGTCTCTGACGCGCATCTCGGCGCCGAG 60

QY 61 AGCTTTTATAGGTTGCTTTTCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120

DB 61 AGCTTTTATAGGTTGCTTTTCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120

QY 121 CCCATATCATCAAGGAGCTGATAATCTAGTGAAGAGTTAGACTGTGCATCTTCACTA 180

DB 121 CCCATATCATCAAGGAGCTGATAATCTAGTGAAGAGTTAGACTGTGCATCTTCACTA 180

QY 181 TGATATGAGGCGAGTCTCTGAGCTTATTTCTCTGTGGAAGATGTGACATATCCAGCGGA 240

DB 181 TGATATGAGGCGAGTCTCTGAGCTTATTTCTCTGTGGAAGATGTGACATATCCAGCGGA 240

QY 241 ACATCATGATGAGGGAACACATGTACAGATGTCTCCACCGGACGAGGCGTC 300

DB 241 ACATCATGATGAGGGAACACATGTACAGATGTCTCCACCGGACGAGGCGTC 300

QY 301 CCGGAGGACGAGGAGCATGTGAGCCAGACCTCTCAGCACATCTTTTAGGCCCCAAAATC 360

DB 301 CCGGAGGACGAGGAGCATGTGAGCCAGACCTCTCAGCACATCTTTTAGGCCCCAAAATC 360

QY 361 TGAGGCTGCTTCACTCCCTCAGCAGCTCTGTGCAATATCAATATGAACTTCAAGTGCCC 420

Db 361 TGAGGCTGCTTACCCCTCAGCAGCTCCTGTGCAATATCAATATGAACCTCCAAGTGCCC 420
Qy 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCTCCACGACCAAGCTTTG 480
Db 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCTCCACGACCAAGCTTTG 480
Qy 481 TACCCCTTCCCCACCAAGCTCCTCCGTGAGCGCAAGCGCCCTCTTCCGCCCTGCCCCAATCA 540
Db 481 TACCCCTTCCCCACCAAGCTCCTCCGTGAGCGCAAGCGCCCTCTTCCGCCCTGCCCCAATCA 540
Qy 541 GCGCGCTTTTCCCAACCAACAGATGAGGCAAGCGCCCTTCCAGTTCTCCTTGTGTTTCCCTC 600
Db 541 GCGCGCTTTTCCCAACCAACAGATGAGGCAAGCGCCCTTCCAGTTCTCCTTGTGTTTCCCTC 600
Qy 601 CCAATGCCACCAATAGCTTGTCTTAATAAAGCGGAGGCTCCTGAGGCAAGCTTCTGAGC 660
Db 601 CCAATGCCACCAATAGCTTGTCTTAATAAAGCGGAGGCTCCTGAGGCAAGCTTCTGAGC 660
Qy 661 AAGGCACTTTCCCTTTCATGATGCCCTCCTCCATGCTCATCCCGCCGCCCTCCAG 720
Db 661 AAGGCACTTTCCCTTTCATGATGCCCTCCTCCATGCTCATCCCGCCGCCCTCCAG 720
Qy 721 TCATGCCGAGCAGGTTAATTAATCAGTACCTCCGGGCTAATCTCACCAACACTTCCAC 780
Db 721 TCATGCCGAGCAGGTTAATTAATCAGTACCTCCGGGCTAATCTCACCAACACTTCCAC 780
Qy 781 CTCGCCAGTTTAAATAGTTTCCAGAACCAAGCTAGTTTCTTCTGCCCCAGTCTAATAACA 840
Db 781 CTCGCCAGTTTAAATAGTTTCCAGAACCAAGCTAGTTTCTTCTGCCCCAGTCTAATAACA 840
Qy 841 GCAGTAGTCTCATTTTTCAGACATCTCCCTCCATACCACATCCCAAGGCTCCAGTGAGA 900
Db 841 GCAGTAGTCTCATTTTTCAGACATCTCCCTCCATACCACATCCCAAGGCTCCAGTGAGA 900
Qy 901 GAAAGTCTCCAGAAAGGCTGAAACACTATGATGACCAAGGACCAAGGACCAAGTCAAG 960
Db 901 GAAAGTCTCCAGAAAGGCTGAAACACTATGATGACCAAGGACCAAGGACCAAGTCAAG 960
Qy 961 GCGAGGTGAGAGGCTCGTCCCTGGATCGGCGGAGGAGCGCGAGTCCGACAGGA 1020
Db 961 GCGAGGTGAGAGGCTCGTCCCTGGATCGGCGGAGGAGCGCGAGTCCGACAGGA 1020
Qy 1021 GAAGACAGACAGCCGGTACAGATCTGATATGACCGAGGAGAACCAATCTCGCCACC 1080
Db 1021 GAAGACAGACAGCCGGTACAGATCTGATATGACCGAGGAGAACCAATCTCGCCACC 1080
Qy 1081 GCAGCTTACGAACGAGCAGAGCAGAGAACCGGAGAGACACAGGCAATCAGACCAACGAA 1140
Db 1081 GCAGCTTACGAACGAGCAGAGCAGAGAACCGGAGAGACACAGGCAATCAGACCAACGAA 1140
Qy 1141 GATCACCCTCTCTGAAAGGTCCTACAAAAAGAGTATAAGAGATCTGGAAGAGTTACG 1200
Db 1141 GATCACCCTCTCTGAAAGGTCCTACAAAAAGAGTATAAGAGATCTGGAAGAGTTACG 1200
Qy 1201 GTTTATCGGTGTTCTGAACTCTGATGATGACACAGATTAACCTGCGGAGATTAATTA 1260
Db 1201 GTTTATCGGTGTTCTGAACTCTGATGATGACACAGATTAACCTGCGGAGATTAATTA 1260
Qy 1261 AAAATACAGATTTCTGGGCCCCACCCCTGGAGATGATGATCATCCCTCCCAAGTAGG 1320
Db 1261 AAAATACAGATTTCTGGGCCCCACCCCTGGAGATGATGATCATCCCTCCCAAGTAGG 1320
Qy 1321 AGAAGAGAGAGCTCTGTTGGAGAGAGAAAAAGACCGTTGGAGTGACAAACAGAGTTCTG 1380
Db 1321 AGAAGAGAGAGCTCTGTTGGAGAGAGAAAAAGACCGTTGGAGTGACAAACAGAGTTCTG 1380
Qy 1381 GCAAGACAGAACTATATCTCAATCAAGGAAAAAGAGCCCGGAGGAGACATGCTGAC 1440
Db 1381 GCAAGACAGAACTATATCTCAATCAAGGAAAAAGAGCCCGGAGGAGACATGCTGAC 1440
Qy 1441 AGAATCAGGAGGAGAGAGACTTCTTAAGCCTGTTGATTCGATGCATCTCAATTCAG 1500
Db 1441 AGAATCAGGAGGAGAGAGAACTTCTTAAGCCTGTTGATTCGATGCATCTCAATTCAG 1500

Qy 1501 AAAAATACTACTCCAGTGAACCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA 1560
Db 1501 AAAAATACTACTCCAGTGAACCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA 1560
Qy 1561 GTAGGCTTCTGCTGACTTATATGACAAATTTGAGGAGGATTGGGAGCAGCAAGAAAGG 1620
Db 1561 GTAGGCTTCTGCTGACTTATATGACAAATTTGAGGAGGATTGGGAGCAGCAAGAAAGG 1620
Qy 1621 CCAAGCTGCTCGGCCCTCGTGGGAACTCCAAAGACGAAGCTCGATGAAGATTAGAGA 1680
Db 1621 CCAAGCTGCTCGGCCCTCGTGGGAACTCCAAAGACGAAGCTCGATGAAGATTAGAGA 1680
Qy 1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACTGTTCTAGCAGTCAAGACT 1740
Db 1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACTGTTCTAGCAGTCAAGACT 1740
Qy 1741 CTGAAGTTTTTCAGCGTTATTCAGAGAAATCAAAAGCGAAAAAGGCCACCTGACCGACTTC 1800
Db 1741 CTGAAGTTTTTCAGCGTTATTCAGAGAAATCAAAAGCGAAAAAGGCCACCTGACCGACTTC 1800
Qy 1801 ATGATGAACCTTTGGTACAAAGATCCAGGCGAGATGAATGATGACCACTCTGCAATGCA 1860
Db 1801 ATGATGAACCTTTGGTACAAAGATCCAGGCGAGATGAATGATGACCACTCTGCAATGCA 1860
Qy 1861 GGCBAAGGCAAGACGACAGGAAATGAGGACAGCAATTTATCTGAGGAGAGGCCATCA 1920
Db 1861 GGCBAAGGCAAGACGACAGGAAATGAGGACAGCAATTTATCTGAGGAGAGGCCATCA 1920
Qy 1921 AGCCCTGCTGCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCAGAGTCT 1980
Db 1921 AGCCCTGCTGCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCAGAGTCT 1980
Qy 1981 CCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAATACGATGATCAGAGT 2040
Db 1981 CCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAATACGATGATCAGAGT 2040
Qy 2041 ATATCTTTGAAAGGATTTCTATGTTTGCACATGCCCCCTCGACCAATATTTCACTGTGTA 2100
Db 2041 ATATCTTTGAAAGGATTTCTATGTTTGCACATGCCCCCTCGACCAATATTTCACTGTGTA 2100
Qy 2101 AAGTAATAGATTCAACATAGACTACCGATTTCATTTCAATGAAGAGATGATGCCGGAGA 2160
Db 2101 AAGTAATAGATTCAACATAGACTACCGATTTCATTTCAATGAAGAGATGATGCCGGAGA 2160
Qy 2161 ATTTTGTGTGAAAGGCTTGAACCTTTTCACTGTGTTCTATTCAGAGATATTTTGAAT 2220
Db 2161 ATTTTGTGTGAAAGGCTTGAACCTTTTCACTGTGTTCTATTCAGAGATATTTTGAAT 2220
Qy 2221 TATATGACTGGAATCTTAAAGGTCCTTTGTTTGAAGACAGCCCTCCCTGCTGCCAAGAT 2280
Db 2221 TATATGACTGGAATCTTAAAGGTCCTTTGTTTGAAGACAGCCCTCCCTGCTGCCAAGAT 2280
Qy 2281 TTTCAATTCATGCGAGTGTGTAAGATTTCTTCCAGATGCGAGGAGAGGAGTCTCTCCA 2340
Db 2281 TTTCAATTCATGCGAGTGTGTAAGATTTCTTCCAGATGCGAGGAGAGGAGTCTCTCCA 2340
Qy 2341 TGCACACAGATTTCTCTGACTTGTGTTAAGGTGAGCAAAAGCCCTGCTGAGGAGGAGA 2400
Db 2341 TGCACACAGATTTCTCTGACTTGTGTTAAGGTGAGCAAAAGCCCTGCTGAGGAGGAGA 2400
Qy 2401 TTGCCAAATATGCTTCAAGTGGGAGGAGTGGAGTGGCAGAAATATGAGAGAAATGCAAG 2460
Db 2401 TTGCCAAATATGCTTCAAGTGGGAGGAGTGGAGTGGCAGAAATATGAGAGAAATGCAAG 2460
Qy 2461 GCATGATGTTTACCAACCTCGGAGCAACCAAGCTCTGTCGGTATCGATCAACTGGATC 2520
Db 2461 GCATGATGTTTACCAACCTCGGAGCAACCAAGCTCTGTCGGTATCGATCAACTGGATC 2520
Qy 2521 GTGAACAGTTCAACCCCGATGATTAATTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580
Db 2521 GTGAACAGTTCAACCCCGATGATTAATTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580

Db 4741 TAAAAAAAAAAAAAAAAAAAAA 4764

RESULT 3
AEB47410
ID AEB47410 standard; cDNA; 4764 BP.
XX
AC AEB47410;
XX
DT 22-SEP-2005 (first entry)
DE Human ribonuclease III (RNase III) cDNA, SEQ ID NO: 1.
XX
XX Ribonuclease; gene silencing; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 246..4370
XX /tag= a
XX /product= "Human ribonuclease III (RNase III) protein"
XX /transl_except= (pos:297..299, aa:Cys)
XX /transl_except= (pos:3876..3879, aa:Gly)
XX /transl_except= (pos:4278..4280, aa:Gly)
XX
XX US2005159384-A1.
XX
XX 21-JUL-2005.
XX
XX 02-DEC-2004; 2004US-00001993.
XX
XX 06-JUL-2001; 2001US-00900425.
XX 20-FEB-2002; 2002US-00079185.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Crooke ST;
XX
XX WPI; 2005-512270/52.
XX P-PSDB; AEB47411.
XX GENBANK; AF189011.
XX
XX Use of RNase III for eliciting modification of a selected RNA target,
XX PT promoting gene silencing of a gene, inhibiting the expression of a gene,
XX PT promoting inhibition of expression of a gene, or eliciting modification
XX PT of an RNA target.
XX
XX Example 2; SEQ ID NO 1; 32pp; English.
XX
XX The present invention relates to ribonuclease III (RNase III) nucleic
XX acids and their encoding proteins. RNase III is an endoribonuclease that
XX cleaves double stranded RNA. The invention is useful for eliciting
XX modification of a selected RNA target in a cell, promoting gene silencing
XX of a gene, inhibiting the expression of a gene, promoting inhibition of
XX expression of a gene and eliciting modification of a RNA target in a
XX cell. The present sequence is human ribonuclease III (RNase III) cDNA.
XX
XX Sequence 4764 BP; 1379 A; 1140 C; 1082 G; 1163 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 4764; DB 14; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCTTGTGACCTGGGTAGTACCTGGCTTGTCTGACGCGGATCTCGGGCCCGAG 60
Db 1 CTGCTTGTGACCTGGGTAGTACCTGGCTTGTCTGACGCGGATCTCGGGCCCGAG 60
QY 61 AGCCTTTTATAGTGTGCTTTTCCCGGGGATGTGAAGATACAGAAATGACTGTGAATCAA 120
Db 61 AGCCTTTTATAGTGTGCTTTTCCCGGGGATGTGAAGATACAGAAATGACTGTGAATCAA 120
QY 121 CCATATCATCAAGGAGCTGATTAATCTAGTGAAGAGTTAGAGTGTGCATCTTCACTA 180
Db 121 GTTTATCGGTTGTCTCTGAACCTGTGGATGCAACCAAGATTACCTGGGGAGATTATTA 1260

121 CCATATCATCAAGGAGCTGATAA TCTAGTGGAGAGTTAGAGCTGTGCATCTTCACTA 180
181 TGATATGAGGAGCTCTCTGAGCTTATATTTCTGTGTGAAGATGTGATATCCAGCGGA 240
181 TGATATGAGGAGCTCTCTGAGCTTATATTTCTGTGTGAAGATGTGATATCCAGCGGA 240
241 ACATCATGATGAGGAGAAACATGTCACAGAAATGTCGTTCCACCCGGGACGAGGGGTC 300
241 ACATCATGATGAGGAGAAACATGTCACAGAAATGTCGTTCCACCCGGGACGAGGGGTC 300
301 CCGGAGGAGGAGGACATGGAGCCAGACCCCTCAGCACCATCTCTTTAGGCCCAAAATC 360
301 CCGGAGGAGGAGGACATGGAGCCAGACCCCTCAGCACCATCTCTTTAGGCCCAAAATC 360
361 TGAGGCTGTTCACTCCCTCAGCGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
361 TGAGGCTGTTCACTCCCTCAGCGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
421 CTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACGACTTTG 480
421 CTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACGACTTTG 480
481 TACCTTCCCTCCACCATGCTCCGTGAGCGCAAGGCCCTCTTCCCTCCCTGCCCATAACA 540
481 TACCTTCCCTCCACCATGCTCCGTGAGCGCAAGGCCCTCTTCCCTCCCTGCCCATAACA 540
541 GGCCTCTTTCCCAACCAACAGAGAGGACCCCTTCCAGTTCCTCTCTGTTTCTCTC 600
541 GGCCTCTTTCCCAACCAACAGAGAGGACCCCTTCCAGTTCCTCTCTGTTTCTCTC 600
601 CCATGCCACCAATGCTTGTCTTAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
601 CCATGCCACCAATGCTTGTCTTAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
661 AAGGCACTTTTCCCTCTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
661 AAGGCACTTTTCCCTCTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
721 TCATGCCAGCAGGTTAATATATCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
721 TCATGCCAGCAGGTTAATATATCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
781 CTCCTCCAGTTTAAATAGTTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
781 CTCCTCCAGTTTAAATAGTTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
841 GCAGTAGTCTCTATTTTCAAGATCTCCCTCCATACCACTCCCAAGGCTCCAGTGAGA 900
841 GCAGTAGTCTCTATTTTCAAGATCTCCCTCCATACCACTCCCAAGGCTCCAGTGAGA 900
901 GAAGGTCCCAGAGAGGCTGAAACACTATGATGACACACAGGAGGAGGAGGAGGAGGAGGAG 960
901 GAAGGTCCCAGAGAGGCTGAAACACTATGATGACACACAGGAGGAGGAGGAGGAGGAGGAG 960
961 GCGAGGTGAGAGGATCGGTCCTGGATCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
961 GCGAGGTGAGAGGATCGGTCCTGGATCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1021 GAAGCAACAGACAGCGGTACAGATCTGATTATGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
1021 GAAGCAACAGACAGCGGTACAGATCTGATTATGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
1081 GCAGCTTACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
1081 GCAGCTTACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
1141 GATCAACATCTCTGGAAGGCTCTCAAAAAAGAGATATGAGAGATCTGGAAGGAGGTACG 1200
1141 GATCAACATCTCTGGAAGGCTCTCAAAAAAGAGATATGAGAGATCTGGAAGGAGGTACG 1200
1201 GTTTATCGGTTCTTCTGAACTGTGGATGCAACCAAGATTACCTGGGGAGATTATTA 1260
1201 GTTTATCGGTTCTTCTGAACTGTGGATGCAACCAAGATTACCTGGGGAGATTATTA 1260

QY 1261 AAAATACAGATCTTTGGGCCCCCACCCTCGAGATTGTGAATCATGCTCCCCAAGTAGGG 1320
DB 1261 AAAATACAGATCTTTGGGCCCCCACCCTCGAGATTGTGAATCATGCTCCCCAAGTAGGG 1320
QY 1321 AGAAGAGAGAGCTCGTTGGGAGGAGAAAAAGACCGTTGGAGTGACAAACAGAGTTCTG 1380
DB 1321 AGAAGAGAGAGCTCGTTGGGAGGAGAAAAAGACCGTTGGAGTGACAAACAGAGTTCTG 1380
QY 1381 GCAAGACAAGAACTATATCACTCAATCAAGGAAAAAGAGCCGCGAGGAGACCATGCTGACA 1440
DB 1381 GCAAGACAAGAACTATATCACTCAATCAAGGAAAAAGAGCCGCGAGGAGACCATGCTGACA 1440
QY 1441 AGAATGAGAGGAGAGAGAGAACTCTTTAAGCCTGTGTGGATTCGATGCATCTATTTCAG 1500
DB 1441 AGAATGAGAGGAGAGAGAGAACTCTTTAAGCCTGTGTGGATTCGATGCATCTATTTCAG 1500
QY 1501 AAAAATACTACTCCAGTGACCCCATCGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA 1560
DB 1501 AAAAATACTACTCCAGTGACCCCATCGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA 1560
QY 1561 GTAGGCTTCGTGACTTATATGACAAATTTGAGAGGAGTTGGGAGCAGGCAAGAAAAAGG 1620
DB 1561 GTAGGCTTCGTGACTTATATGACAAATTTGAGAGGAGTTGGGAGCAGGCAAGAAAAAGG 1620
QY 1621 CCNAAAGCTCTCGGCCCTCGTGGGAACCTCCNAAAGCAGAGCTCGATGAAGATTTAGAGA 1680
DB 1621 CCNAAAGCTCTCGGCCCTCGTGGGAACCTCCNAAAGCAGAGCTCGATGAAGATTTAGAGA 1680
QY 1681 GTTCCAGTCAATCCGAGTGTGATCGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1740
DB 1681 GTTCCAGTCAATCCGAGTGTGATCGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1740
QY 1741 CTGAAGTTTTTGAAGTTATGACGAAATCAAAACGCAAAAAAGGCCACCTGACCGACTTC 1800
DB 1741 CTGAAGTTTTTGAAGTTATGACGAAATCAAAACGCAAAAAAGGCCACCTGACCGACTTC 1800
QY 1801 ATGATGAACTTTGGTACAGATCCAGGCCAGTGAATGATGACCACTCTGCNAAATGCA 1860
DB 1801 ATGATGAACTTTGGTACAGATCCAGGCCAGTGAATGATGACCACTCTGCNAAATGCA 1860
QY 1861 GGCNAAAGCAGACGCAAGGAAATTAGGCACAGCAATTTATCTGTGAGAGAGGCCATCA 1920
DB 1861 GGCNAAAGCAGACGCAAGGAAATTAGGCACAGCAATTTATCTGTGAGAGAGGCCATCA 1920
QY 1921 AGCCCTGTGCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCAGACTCT 1980
DB 1921 AGCCCTGTGCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCAGACTCT 1980
QY 1981 CCCCCTACGAACTTTTAACTGACAGCCCACTGTTATAGNATACGATGATCAGAGT 2040
DB 1981 CCCCCTACGAACTTTTAACTGACAGCCCACTGTTATAGNATACGATGATCAGAGT 2040
QY 2041 ATATCTTTGAAGGATTTCTATGTTGCAATGCCCCCTGACCAATATTTCCACTGTGTA 2100
DB 2041 ATATCTTTGAAGGATTTCTATGTTGCAATGCCCCCTGACCAATATTTCCACTGTGTA 2100
QY 2101 AAGTAAATAGATTTCAACATAGACTACACGATTCATTTCAATGAGAGATGATGCGGAGA 2160
DB 2101 AAGTAAATAGATTTCAACATAGACTACACGATTCATTTCAATGAGAGATGATGCGGAGA 2160
QY 2161 ATTTTGTGTGAAGGCTTGAACTCTTTTCACTGTTCTTATTCAGAGATATTTTGAAT 2220
DB 2161 ATTTTGTGTGAAGGCTTGAACTCTTTTCACTGTTCTTATTCAGAGATATTTTGAAT 2220
QY 2221 TATATGACTGGAATCTTTAAGGTCTTTGTTTCAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
DB 2221 TATATGACTGGAATCTTTAAGGTCTTTGTTTGAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
QY 2281 TTCAATTTCAATGCAAGCTTTTGTGAAGATTTCTCCAGATGAGGAAAGGAAAGTCTCTCCA 2340
DB 2281 TTCAATTTCAATGCAAGCTTTTGTGAAGATTTCTCCAGATGAGGAAAGGAAAGTCTCTCCA 2340

QY 2341 TGACACAGATTCCTCTGTACTTTTAAAGTGACAGAAAGCCCTGTGTGCTGAGGAGAGA 2400
DB 2341 TGACACAGATTCCTCTGTACTTTTAAAGTGACAGAAAGCCCTGTGTGCTGAGGAGAGA 2400
QY 2401 TTGCCAAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGACAGNAGATGCAAG 2460
DB 2401 TTGCCAAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGACAGNAGATGCAAG 2460
QY 2461 GCATGATTTGTACCAACCTCGGAGCAAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
DB 2461 GCATGATTTGTACCAACCTCGGAGCAAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
QY 2521 GTGAACAGTTCAACCCCGATGTGATTTATCTTTCCGATATCGTCCACTTTGGGATACGCC 2580
DB 2521 GTGAACAGTTCAACCCCGATGTGATTTATCTTTCCGATATCGTCCACTTTGGGATACGCC 2580
QY 2581 CTGCAAGTTGAGTTATGACGAGAGACCCACAGTACCAAAACTGTGGAAGAGTTATGTGA 2640
DB 2581 CTGCAAGTTGAGTTATGACGAGAGACCCACAGTACCAAAACTGTGGAAGAGTTATGTGA 2640
QY 2641 AACTTGGCCACCTCTAGCAAAATAGTCCCAAAAGTCAAACTGACAAACAGAGCTGG 2700
DB 2641 AACTTGGCCACCTCTAGCAAAATAGTCCCAAAAGTCAAACTGACAAACAGAGCTGG 2700
QY 2701 CACAGAGGAGGAGGAGCCCTCCAAAAATAACGACAGAGAAATCAATGACACAGAGTAA 2760
DB 2701 CACAGAGGAGGAGGAGCCCTCCAAAAATAACGACAGAGAAATCAATGACACAGAGTAA 2760
QY 2761 CGGTGAGCTAAGTAGCCAAAGATTCTGGAAAACTGGGCATCCGTTCTGTGTCAGC 2820
DB 2761 CGGTGAGCTAAGTAGCCAAAGATTCTGGAAAACTGGGCATCCGTTCTGTGTCAGC 2820
QY 2821 ATGCAATGATGCTGTTCTGACCCATCATATCCGCTACCACTAATGCCCTAATGCAAT 2880
DB 2821 ATGCAATGATGCTGTTCTGACCCATCATATCCGCTACCACTAATGCCCTAATGCAAT 2880
QY 2881 TGGACAAGTTGATAGGATATATCTTTCCAAAGATCGTTGTCTGTGACGTGGCCATGACTC 2940
DB 2881 TGGACAAGTTGATAGGATATATCTTTCCAAAGATCGTTGTCTGTGACGTGGCCATGACTC 2940
QY 2941 ATCCAAGTCATCTTTAAATTTTGAATGAAATCTGATCATCGCAGGAATTCATTTATCTA 3000
DB 2941 ATCCAAGTCATCTTTAAATTTTGAATGAAATCTGATCATCGCAGGAATTCATTTATCTA 3000
QY 3001 ACTGTGGAATTCGGCAGCCCAATACGAGACAGAAAAAGTTCAATCATGACATCGCGA 3060
DB 3001 ACTGTGGAATTCGGCAGCCCAATACGAGACAGAAAAAGTTCAATCATGACATCGCGA 3060
QY 3061 AGNAAAGGATTAACACCTTTGATAAATATCATGTCAGCCCTTGGCCAAGATGACCCAACTC 3120
DB 3061 AGNAAAGGATTAACACCTTTGATAAATATCATGTCAGCCCTTGGCCAAGATGACCCAACTC 3120
QY 3121 CTTGAGGATTAACCAATGAACGGTTGGAATTTCTGGGTGATGCTGTTGTTGAATTTTC 3180
DB 3121 CTTGAGGATTAACCAATGAACGGTTGGAATTTCTGGGTGATGCTGTTGTTGAATTTTC 3180
QY 3181 TGAACAGCTCTCATTTGTACTATTTGTTTCTAGTCTGGAAGAAGAGGATTAGCAACT 3240
DB 3181 TGAACAGCTCTCATTTGTACTATTTGTTTCTAGTCTGGAAGAAGAGGATTAGCAACT 3240
QY 3241 ATCGGACTGCCATTTGTTAGAAATCAGACCTTTGCCATGCTAGCAAAAGAACTTGAACCTGG 3300
DB 3241 ATCGGACTGCCATTTGTTAGAAATCAGACCTTTGCCATGCTAGCAAAAGAACTTGAACCTGG 3300
QY 3301 ATCCATTTATGCTGATGCTCAGCGGCTTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
DB 3301 ATCCATTTATGCTGATGCTCAGCGGCTTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
QY 3361 CAATGGCCAAATTTGTTTGAAGCTTTTAAATAGGAGCTGTTTACTTTGAGGGAAGCTCGAGG 3420
DB 3361 CAATGGCCAAATTTGTTTGAAGCTTTTAAATAGGAGCTGTTTACTTTGAGGGAAGCTCGAGG 3420
QY 3421 AAGCCAAAGCAGTTATTTTGAAGCTTCTCTTTAAATGATCCGGAACCTGCGCGAAGTCTGGC 3480

Db 3421 AAGCCAGAGATTATTTGGAGCTTGCTCTTAAATGATCGGACCTGCGGAGTCTGCG 3480
Qy TCAATATATCTCTCCACCACTCCAACTACAAGAGCCAAATCTGATCGACAATTTATTG 3540
Db TCAATATATCTCTCCACCACTCCAACTACAAGAGCCAAATCTGATCGACAATTTATTG 3540
Qy AAACCTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAAGCAATGGAGTAATTTTAA 3600
Db AAACCTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAAGCAATGGAGTAATTTTAA 3600
Qy CTATCTCTCGACTCTGCGAAGGCAATTCACATTTGAGAACTGTGGGATTTAAACATCTGA 3660
Db CTATCTCTCGACTCTGCGAAGGCAATTCACATTTGAGAACTGTGGGATTTAAACATCTGA 3660
Qy CCTAGGCCACAATCAGAGAATGGAATCTTAGTGTGACTCCATAATGCAACTGGTAGCCA 3720
Db CCTAGGCCACAATCAGAGAATGGAATCTTAGTGTGACTCCATAATGCAACTGGTAGCCA 3720
Qy CAGAGTACTTATTCATTTCCAGATCATCATGAGACACTTAACTTTGTTGCGAA 3780
Db CAGAGTACTTATTCATTTCCAGATCATCATGAGACACTTAACTTTGTTGCGAA 3780
Qy GCTCTTTGTTGTAATATAGAACTCAGGCCAAGGTAGCGAGGCTGGGCATGCGAGCT 3840
Db GCTCTTTGTTGTAATATAGAACTCAGGCCAAGGTAGCGAGGCTGGGCATGCGAGCT 3840
Qy AGCCATAACCAACGACAAGAGCGCTGTGGCGCTTCGCCACCAAGACTTTGGCGG 3900
Db AGCCATAACCAACGACAAGAGCGCTGTGGCGCTTCGCCACCAAGACTTTGGCGG 3900
Qy ACCTTTGGAATCATTTATTCAGCGCTGTACATGATGAAGATTGGGAATATGTTTATA 3960
Db ACCTTTGGAATCATTTATTCAGCGCTGTACATGATGAAGATTGGGAATATGTTTATA 3960
Qy CTTTCATGATGTCGCTCTTTCCAGCAATGGAAGAAATTCATTTTGAATCAGGATTGA 4020
Db CTTTCATGATGTCGCTCTTTCCAGCAATGGAAGAAATTCATTTTGAATCAGGATTGA 4020
Qy ATGACCCCAATCCAGCTTCAGCAGTGTCTTGACACTTAGGACAGAGGAAAGAGC 4080
Db ATGACCCCAATCCAGCTTCAGCAGTGTCTTGACACTTAGGACAGAGGAAAGAGC 4080
Qy CAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCCCATCCCATGCCGAACTTACA 4140
Db CAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCCCATCCCATGCCGAACTTACA 4140
Qy CTGTGGCTGTTTATTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Db CTGTGGCTGTTTATTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Qy AAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAAAAATATAATTTTCCCGAGATGGCCC 4260
Db AAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAAAAATATAATTTTCCCGAGATGGCCC 4260
Qy ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTAAAGAAATGAGGTGGGAA 4320
Db ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTAAAGAAATGAGGTGGGAA 4320
Qy GAGAGCATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATTAAGAGGGCAT 4380
Db GAGAGCATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATTAAGAGGGCAT 4380
Qy GCAAGTGTGGAGTATTACTTGTCTAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
Db GCAAGTGTGGAGTATTACTTGTCTAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
Qy AGTTTCTCGACAAATGAACGAGTGTGCTCATTTGAAATATAATACAGAGTCAATCG 4500
Db AGTTTCTCGACAAATGAACGAGTGTGCTCATTTGAAATATAATACAGAGTCAATCG 4500
Qy CTATGTTGTTTAAATGATCTGTTTACGTGATGCTTTTATTACAAAGTATTAGATT 4560
|||||

Db 4501 CTATTGTTGTTTAAATGATCTGTTTACGTGATGCTCTTATTATTACAAAGTATTAGATT 4560
Qy TTTCTTCTTAAATTAACGGAACCTTGACTTTTGGTGAATGTCATTACTCTCTTTTATTG 4620
Db TTTCTTCTTAAATTAACGGAACCTTGACTTTTGGTGAATGTCATTACTCTCTTTTATTG 4620
Qy TTTCTTAAATATAAAATTCAGAAAGCATATTCTATGTGGAATAGATCCTGTTTTCAT 4680
Db TTTCTTAAATATAAAATTCAGAAAGCATATTCTATGTGGAATAGATCCTGTTTTCAT 4680
Qy CTGTGCTCCAGATTTGTGACCTAGACTTTCAATTTGCAAGTAAATAATTTGACTTTACTAG 4740
Db CTGTGCTCCAGATTTGTGACCTAGACTTTCAATTTGCAAGTAAATAATTTGACTTTACTAG 4740
Qy TTTAAATAATAAAATTTCAAGAAAGCATATTCTATGTGGAATAGATCCTGTTTTCAT 4764
Db TTTAAATAATAAAATTTCAAGAAAGCATATTCTATGTGGAATAGATCCTGTTTTCAT 4764
|||||

RESULT 4
ADG32900
ID ADG32900 standard; DNA; 4764 BP.
XX
AC ADG32900;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human DNA differentially expressed in patients with SLE SeqID224.
XX
KW human; ds; autoimmune; chronic inflammatory disease; SLE;
KW systemic lupus erythematosus; rheumatoid arthritis; cholecystitis;
KW Sjogren's disease; CREST syndrome; scleroderma; ankylosing spondylitis;
KW ulcerative colitis; primary sclerosing cholangitis; appendicitis;
KW diverticulitis; primary biliary sclerosis.
XX
OS Homo sapiens.
XX
PN WO2003090694-A2.
XX
PD 06-NOV-2003.
XX
PF 24-APR-2003; 2003WO-US013015.
XX
PR 24-APR-2002; 2002US-00131827.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N;
XX WPI; 2003-877243/81.
XX
PT Diagnosing or monitoring autoimmune and chronic inflammatory diseases,
PT such as rheumatoid arthritis, systemic lupus erythematosus, ulcerative
PT colitis, psoriasis and asthma by detecting the expression level of one or
PT more genes.
XX
PS Claim 18; SEQ ID NO 224; 877pp; English.
XX
CC This invention relates to novel methods for diagnosing and monitoring
CC autoimmune and chronic inflammatory diseases. Specifically, it refers to
CC the identification of genes that have a clinical utility as diagnostic
CC tools for the management of, in particular, patients with systemic lupus
CC erythematosus (SLE) or rheumatoid arthritis (RA). Accordingly, the
CC present invention describes a method for determining the levels of
CC multiple differentially expressed genes of a patient, in a concerted
CC manner, in order to achieve an improved diagnostic assay with sensitivity
CC and specificity for the disease in question. As such, these genes are
CC useful for the diagnosis of various other inflammatory disorders
CC including cholecystitis, Sjogren's disease, CREST syndrome, scleroderma,
CC ankylosing spondylitis, ulcerative colitis, primary sclerosing
CC cholangitis, appendicitis, diverticulitis, and primary biliary sclerosis.
CC This polynucleotide is a DNA sequence representing human mRNA that is
CC differentially expressed in patients with SLE, used in an exemplification

CC	of the invention.									
XX										
SQ	Sequence 4764 BP; 1379 A; 1138 C; 1082 G; 1165 T; 0 U; 0 Other;									
	Query Match 99.9%; Score 4760.8; DB 10; Length 4764;									
	Best Local Similarity 100.0%; Pred. No. 0;									
	Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CTGTCTTGTA	CTCTGCGGTAGT	AGCTGGCTTTGCTCTGACGGCGATCTCGGGGCCGAG	60					
DB	1	CTGTCTTGTA	CTCTGCGGTAGT	AGCTGGCTTTGCTCTGACGGCGATCTCGGGGCCGAG	60					
QY	61	AGCCTTTTAT	AGTTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	120						
DB	61	AGCCTTTTAT	AGTTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	120						
QY	121	CCCATATCAT	CAAGGAGCTGATAATCTAGTGAAGATTTAGACGTGTGCATCTCACTA	180						
DB	121	CCCATATCAT	CAAGGAGCTGATAATCTAGTGAAGATTTAGACGTGTGCATCTCACTA	180						
QY	181	TGATATGAGG	CAGTCTCTGAGCTTATATCTCTGTGGAGATGTGACATATCCAGGCCGA	240						
DB	181	TGATATGAGG	CAGTCTCTGAGCTTATATCTCTGTGGAGATGTGACATATCCAGGCCGA	240						
QY	241	ACATCATGAT	GACGGGAAACACATGTCAAGAAATGTGTTCCACCGGGACGAGGCGTC	300						
DB	241	ACATCATGAT	GACGGGAAACACATGTCAAGAAATGTGTTCCACCGGGACGAGGCGTC	300						
QY	301	CCCGAGGACG	AGGAGGACATGGAGCGAGACCTCTAGCACCATCTTTTAGGCCCCCAAAATC	360						
DB	301	CCCGAGGACG	AGGAGGACATGGAGCGAGACCTCTAGCACCATCTTTTAGGCCCCCAAAATC	360						
QY	361	TGAGGCTGCT	TACCCCTCAGCGACCTCTCTGTGCAATATCAATATGAACCTCCAAATGTCGC	420						
DB	361	TGAGGCTGCT	TACCCCTCAGCGACCTCTCTGTGCAATATCAATATGAACCTCCAAATGTCGC	420						
QY	421	CTTTCCACCA	CTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACACGACGACTTTG	480						
DB	421	CTTTCCACCA	CTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACACGACGACTTTG	480						
QY	481	TACCCCTTCC	CCACCCATGCTCGGTGAGCGAAGGCCCTCTTCCCTCCCTGCCCATAATCA	540						
DB	481	TACCCCTTCC	CCACCCATGCTCGGTGAGCGAAGGCCCTCTTCCCTCCCTGCCCATAATCA	540						
QY	541	GGCGGCTTTT	CCCAACCAACAGATGAGGACCCCTTCCAGTTCCTCTGTTTTCCTC	600						
DB	541	GGCGGCTTTT	CCCAACCAACAGATGAGGACCCCTTCCAGTTCCTCTGTTTTCCTC	600						
QY	601	CCATGCCACA	CCAATGCTTGTCTTAATAACCCCGCAGTCCCTGGGGCACCTCTCTGGAC	660						
DB	601	CCATGCCACA	CCAATGCTTGTCTTAATAACCCCGCAGTCCCTGGGGCACCTCTCTGGAC	660						
QY	661	AAGGCACTTT	CCCTTCATGATGCCCTCCCTCCATGCTCATCCCCCGCCCTCCAG	720						
DB	661	AAGGCACTTT	CCCTTCATGATGCCCTCCCTCCATGCTCATCCCCCGCCCTCCAG	720						
QY	721	TGATGCCGCA	GCAGGTTAATATACGATCCCTCCGGCTATTCTACACCAACTTCCCA	780						
DB	721	TGATGCCGCA	GCAGGTTAATATACGATCCCTCCGGCTATTCTACACCAACTTCCCA	780						
QY	781	CTCCCAAGTT	TAAATAGTTTCCAGAACACCTCTAGTTCTTTCTGCCCAGTGTCTAATAACA	840						
DB	781	CTCCCAAGTT	TAAATAGTTTCCAGAACACCTCTAGTTCTTTCTGCCCAGTGTCTAATAACA	840						
QY	841	GCAGTAGTCT	CAATTTTACAGACATCTCCCTCCATACCCACTTCCCAAAGGCTCCCAAGTGA	900						
DB	841	GCAGTAGTCT	CAATTTTACAGACATCTCCCTCCATACCCACTTCCCAAAGGCTCCCAAGTGA	900						
QY	901	GAAGTCCCA	GAGGCTGAACACTATGATGACCAACGAGGACCGAGCCACGACTCATG	960						
DB	901	GAAGTCCCA	GAGGCTGAACACTATGATGACCAACGAGGACCGAGATCATGACTCATG	960						
QY	961	GGCGAGGTG	AGGACATCGTCCCTTGGATCGGGGAGCGAGGCCGAGTCCCGACAGGA	1020						

Db 2041 ATATCTTTGAAGGATTTTCTATGTTTGCAATGCCCCCTGACCAATATTTCCAATGTA 2100
Qy 2101 AAGTAATAGATTCAACATAGACTACAGATTCATTTTCATTGGAAGAGATGATGCCGAGA 2160
Db 2101 AAGTAATTAGATTCAACATAGACTACAGATTCATTTTCATTGGAAGAGATGATGCCGAGA 2160
Qy 2161 ATTTTGTGTGAAGGCTTTGAACCTCTTTTTCAGTCTCTTATTCAGAGATATTTTGAAT 2220
Db 2161 ATTTTGTGTGAAGGCTTTGAACCTCTTTTTCAGTCTCTTATTCAGAGATATTTTGAAT 2220
Qy 2221 TATATGACTGGAATCTTTAAAGGTCCTTTGTTGAAGACAGCCCTCGCTGCTGCCAAGAT 2280
Db 2221 TATATGACTGGAATCTTTAAAGGTCCTTTGTTGAAGACAGCCCTCGCTGCTGCCAAGAT 2280
Qy 2281 TTCAATTCATGCCACGTTTGTAAAGATTTCTTCAGATGAGAGAAAGAGATGCTGTCCA 2340
Db 2281 TTCAATTCATGCCACGTTTGTAAAGATTTCTTCAGATGAGAGAAAGAGATGCTGTCCA 2340
Qy 2341 TGACACAGATTTCTCTGTACTTCTTAAGGTGACAGCAAGCCCTGCTGCTGAGGAGAGA 2400
Db 2341 TGACACAGATTTCTCTGTACTTCTTAAGGTGACAGCAAGCCCTGCTGCTGAGGAGAGA 2400
Qy 2401 TTGCCAATATGCTTCAGTGGGAGGCTGGAGTGGCAGAAATATGCAAGAGATGCAAAAG 2460
Db 2401 TTGCCAATATGCTTCAGTGGGAGGCTGGAGTGGCAGAAATATGCAAGAGATGCAAAAG 2460
Qy 2461 GCATGATGTTTACCAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
Db 2461 GCATGATGTTTACCAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
Qy 2521 GTGAAACAGTTCAACCCCGATGTGAATTTCTTTCCGATTTACGTCCTTGGGATACGCC 2580
Db 2521 GTGAAACAGTTCAACCCCGATGTGAATTTCTTTCCGATTTACGTCCTTGGGATACGCC 2580
Qy 2581 CTGCACAGTTGAGTTATGCAAGGAGACCCACAGTACCAGAACTGTGGAAGAGTTATGTGA 2640
Db 2581 CTGCACAGTTGAGTTATGCAAGGAGACCCACAGTACCAGAACTGTGGAAGAGTTATGTGA 2640
Qy 2641 AACTTCGCCACCTCTAGCAATAGTCCCAAGTCAAACTGACAAACAGAGAGCTGG 2700
Db 2641 AACTTCGCCACCTCTAGCAATAGTCCCAAGTCAAACTGACAAACAGAGAGCTGG 2700
Qy 2701 CACAGAGGAGAGAGCCCTCCAAAATACGGCAGAGAAATACAAATGAGACGAGAGTAA 2760
Db 2701 CACAGAGGAGAGAGCCCTCCAAAATACGGCAGAGAAATACAAATGAGACGAGAGTAA 2760
Qy 2761 CGGTGAGCTTAAGTACCAAGGATCTGGAATACTGSCATCCGTTCTGATGCTGTGTCAGC 2820
Db 2761 CGGTGAGCTTAAGTACCAAGGATCTGGAATACTGGAATACTGSCATCCGTTCTGATGCTGTGTCAGC 2820
Qy 2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTACCAATGCTTAATGCATT 2880
Db 2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTACCAATGCTTAATGCATT 2880
Qy 2881 TGGACAAGTTGATAGGATATACTTTCCAGATCGTTGCTGTGTCAGCTGGCCATGACTC 2940
Db 2881 TGGACAAGTTGATAGGATATACTTTCCAGATCGTTGCTGTGTCAGCTGGCCATGACTC 2940
Qy 2941 ATCCAGTCAATCAATTTTAAATTTTGAATGAACTCCGATCATGCGCAGAAATTCATATCTA 3000
Db 2941 ATCCAGTCAATCAATTTTAAATTTTGAATGAACTCCGATCATGCGCAGAAATTCATATCTA 3000
Qy 3001 ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAGTTCAATCAATGCAATGCCGA 3060
Db 3001 ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAGTTCAATCAATGCAATGCCGA 3060
Qy 3061 AGAAAGGATTAACA CTTGTGATAAATATCATGTACGCTTGGCCCAAGATGACCCAACTC 3120
Db 3061 AGAAAGGATTAACA CTTGTGATAAATATCATGTACGCTTGGCCCAAGATGACCCAACTC 3120
Qy 3121 CCTCGAGGATTAACCAATGACGCTGGAATTTCTGGGTGATGCTGTTGTAATTC 3180
Db 3121 CCTCGAGGATTAACCAATGACGCTGGAATTTCTGGGTGATGCTGTTGTAATTC 3180

Qy 3181 TGACCAAGCGTCCATTTTGTACTATTTGTTTCTTAGTCTGGAGAGAGGATTTAGCAACCT 3240
Db 3181 TGACCAAGCGTCCATTTTGTACTATTTTCTTAGTCTGGAGAGAGGATTTAGCAACCT 3240
Qy 3241 ATCGGACTGCAATGTTTCAAGATCAGACCTTCCCATGCTAGCAAGAAAATTTGAACTGG 3300
Db 3241 ATCGGACTGCAATGTTTCAAGATCAGACCTTCCCATGCTAGCAAGAAAATTTGAACTGG 3300
Qy 3301 ATCCATTTATGCTGATGCTCAGCGCCCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
Db 3301 ATCCATTTATGCTGATGCTCAGCGCCCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
Qy 3361 CAATGSCCAATGTTTGTGAAGCGTTAATAGGAGCTGTTTACTTGGAGGAGAGCTTGAGG 3420
Db 3361 CAATGSCCAATGTTTGTGAAGCGTTAATAGGAGCTGTTTACTTGGAGGAGAGCTTGAGG 3420
Qy 3421 AAGCCAAGCAGTTATTTTGAAGCGTCTCTTTAATGATCCGACCTGCGGAAAGTCTGGC 3480
Db 3421 AAGCCAAGCAGTTATTTTGAAGCGTCTCTTTAATGATCCGACCTGCGGAAAGTCTGGC 3480
Qy 3481 TCAATTTATCCTCTCCACCCACTCCAACTACAAGAGCCAAATACTGATCGACAACTTATTC 3540
Db 3481 TCAATTTATCCTCTCCACCCACTCCAACTACAAGAGCCAAATACTGATCGACAACTTATTC 3540
Qy 3541 AAACCTTCTCAGTTCTTCAAAAACTTACTGAGTTTGAAGCAATTTGGAGTAATTTTAA 3600
Db 3541 AAACCTTCTCAGTTCTTCAAAAACTTACTGAGTTTGAAGCAATTTGGAGTAATTTTAA 3600
Qy 3601 CTCATGTTTCGACTTCTGGCAAGGCAATTCACATTTGAGAACTGTGGGATTTTAACTCTGA 3660
Db 3601 CTCATGTTTCGACTTCTGGCAAGGCAATTCACATTTGAGAACTGTGGGATTTTAACTCTGA 3660
Qy 3661 CCTAGGCCACCAATCAGAGAAATGGAATCTCTAGGTGACTCCATAATGCAACTGCTGACCA 3720
Db 3661 CCTAGGCCACCAATCAGAGAAATGGAATCTCTAGGTGACTCCATAATGCAACTGCTGACCA 3720
Qy 3721 CAGAGTACTTATTTCAATTTTCCAGATCATCATGAAGGACACTTAACTTTGTTGCGGAA 3780
Db 3721 CAGAGTACTTATTTCAATTTTCCAGATCATCATGAAGGACACTTAACTTTGTTGCGGAA 3780
Qy 3781 GCTCTTTGTTGTAATATAGAACTCAGGCCAAGGTAGCGGAGGCTGGGCAATGCTGAGAGT 3840
Db 3781 GCTCTTTGTTGTAATATAGAACTCAGGCCAAGGTAGCGGAGGCTGGGCAATGCTGAGAGT 3840
Qy 3841 AGCCCAATACCAACGACAGACCAAGAGGCTGTGGCGCTTCGACCAAGACCTTGGCGG 3900
Db 3841 AGCCCAATACCAACGACAGACCAAGAGGCTGTGGCGCTTCGACCAAGACCTTGGCGG 3900
Qy 3901 ACCTTTTGGAAATCAATTTATGAGCGCTGTACATGATTAAGGATTTGGAAATATGTTCA 3960
Db 3901 ACCTTTTGGAAATCAATTTATGAGCGCTGTACATGATTAAGGATTTGGAAATATGTTCA 3960
Qy 3961 CTTTCAATGATGCTGCTCTTTCCAGATTCGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
Db 3961 CTTTCAATGATGCTGCTCTTTCCAGATTCGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
Qy 4021 ATGACCCCAAAATCCAGCTTTCAGCAGTGTGCTTGACACTTAGGACAGAGGAAAGAGC 4080
Db 4021 ATGACCCCAAAATCCAGCTTTCAGCAGTGTGCTTGACACTTAGGACAGAGGAAAGAGC 4080
Qy 4081 CAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCCCCATCCCATGCCGAACTTACA 4140
Db 4081 CAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCCCCATCCCATGCCGAACTTACA 4140
Qy 4141 CTGTGGCTGTTTATTTCAAGGAGAGAAATAGGCTGTGGGAAAGCAATGTTATTCAGC 4200
Db 4141 CTGTGGCTGTTTATTTCAAGGAGAGAAATAGGCTGTGGGAAAGCAATGTTATTCAGC 4200
Qy 4201 AAGCGGAAATGGGAGCAGCAATGGAATGCGCTTGAATAATATATTTTCCAGATGGCCC 4260
Db 4201 AAGCGGAAATGGGAGCAGCAATGGAATGCGCTTGAATAATATATTTTCCAGATGGCCC 4260

QY 4261 ATCAGAACGGTTTCATCGAACGGAAGTACAGACAGAGTTTAAAGAAATGAGTGGGAAA 4320
DB |||||||
QY 4261 ATCAGAACGGTTTCATCGAACGGAAGTACAGACAGAGTTTAAAGAAATGAGTGGGAAA 4320
DB |||||||
QY 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGAGGGCAT 4380
DB |||||||
QY 4381 GCAAGTGTGGAGTATTACTTGTCTAGTAACTGTGACTGTGCTATTGAGACCTTAGCCT 4440
DB |||||||
QY 4441 AGTTTTCTTCGACAGCAATGAACGAAGTGTGCTCATTTGAATAAAATACAGAGTCAAAATCG 4500
DB |||||||
QY 4501 CTATTGTTGTTTAAATGATCTGTTTTAGCTGGATGGTCTTTTATTACAAAGTATTAGATT 4560
DB |||||||
QY 4561 TTTCTTCTATTAAACGGAAACTTGAATGGAATGTCATTGGAATTAATAACAGAGTCAAAATCG 4620
DB |||||||
QY 4621 CTCCTTTAAATATAAAATTAAGAAAGTATTTAGCTGGATGATCTGTTTTTCCAT 4680
DB |||||||
QY 4681 CTGTGTCAGAGTGTGACCCCTAGACTTTTCAATTGACAAGTAAATAATTTGACTTTTACTAG 4740
DB |||||||
QY 4741 TAAAAAATTAATAAAATTAAGAAAGTATTTAGCTGGATGATCTGTTTTTCCAT 4764
DB |||||||

RESULT 5
ADQ17463
XX ADQ17463 standard; DNA; 4764 BP.
AC ADQ17463;
DT 26-AUG-2004 (first entry)
XX
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 280.
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.
XX Homo sapiens.
OS
XX WO2004048938-A2.
PN
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
DR
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 280; 210pp; English.
XX

CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 4764 BP; 1379 A; 1138 C; 1082 G; 1165 T; 0 U; 0 Other;

Query Match 99.9%; Score 4760.8; DB 12; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGTCTTGCTACCTGCGGTAGTAGCTGGCTTGTCTCTGACGGCGATCTCGGGCCCGAG 60
DB |||||||
QY 1 CTGTCTTGCTACCTGCGGTAGTAGCTGGCTTGTCTCTGACGGCGATCTCGGGCCCGAG 60
DB |||||||
QY 61 AGCCTTTTATAGTGTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
DB |||||||
QY 61 AGCCTTTTATAGTGTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
DB |||||||
QY 121 CCATATATCATCAAGAGCTGATAATCTAGTGAAGATTAGACGTGTGCATCTTCACTA 180
DB |||||||
QY 121 CCATATATCATCAAGAGCTGATAATCTAGTGAAGATTAGACGTGTGCATCTTCACTA 180
DB |||||||
QY 181 TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCGGA 240
DB |||||||
QY 181 TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCGGA 240
DB |||||||
QY 241 ACATCATGATGAGGAGAAACACATGTTCAGAAATGTCTTCCACCCGGGACGAGGCGTC 300
DB |||||||
QY 241 ACATCATGATGAGGAGAAACACATGTTCAGAAATGTCTTCCACCCGGGACGAGGCGTC 300
DB |||||||
QY 301 CCCGAGGACGAGGAGGACATGAGCCAGACCCCTCAGACCATCTCTTTAGGCCCCAAAATC 360
DB |||||||
QY 301 CCCGAGGACGAGGAGGACATGAGCCAGACCCCTCAGACCATCTCTTTAGGCCCCAAAATC 360
DB |||||||
QY 361 TGAGGCTGCTTCACTCCCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
DB |||||||
QY 361 TGAGGCTGCTTCACTCCCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
DB |||||||
QY 421 CTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACGACTTTG 480
DB |||||||
QY 421 CTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACGACTTTG 480
DB |||||||
QY 481 TACCTTTCCCCCAACCCATGCTCCGTGACGGAGAGGCCCTCTTCCCTCCCTGCGCAATCA 540
DB |||||||
QY 481 TACCTTTCCCCCAACCCATGCTCCGTGACGGAGAGGCCCTCTTCCCTCCCTGCGCAATCA 540
DB |||||||
QY 541 GGGCGCTTTTCCCAACCCATGCTCCGTGACGGAGAGGCCCTCTTCCCTCCCTGCGCAATCA 600
DB |||||||
QY 541 GGGCGCTTTTCCCAACCCATGCTCCGTGACGGAGAGGCCCTCTTCCCTCCCTGCGCAATCA 600
DB |||||||
QY 601 CCATGCGACCAATGCTTGTCTTAATAAACCCCTCCAGTCCCTGGGGACCTCTCTGGAC 660
DB |||||||
QY 601 CCATGCGACCAATGCTTGTCTTAATAAACCCCTCCAGTCCCTGGGGACCTCTCTGGAC 660
DB |||||||
QY 661 AAGGCATTTTCCCTTCAATGATGCCCTTCCCTCCATGCTCATCCCCCGCCCCCTCCAG 720
DB |||||||
QY 721 TCATGCGGACGAGGTATTATTCAGTACCTTCCGGCTTATCTCACCACACTTCCAC 780
DB |||||||
QY 721 TCATGCGGACGAGGTATTATTCAGTACCTTCCGGCTTATCTCACCACACTTCCAC 780
DB |||||||
QY 781 CTCCCAGTTTTTAATAGTTTTCCAGAAACACCTTAGTTCTTTTCTGCGCCAGTGTAAATACA 840
DB |||||||

||||| 781 CTCCAGTTTAAATAGTTTCCAGAACACCCTAGTTCTTCTCCAGGCTCTAATAACA 840
QY 841 GCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACATCCCAAAGGCTCCAGTGAGA 900
Db 841 GCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACATCCCAAAGGCTCCAGTGAGA 900
QY 901 GAAGGTCCCCAGAAAGGCTGAACACTATGATACCAACGACGACCCGAGACCAAGTCATG 960
Db 901 GAAGGTCCCCAGAAAGGCTGAACACTATGATACCAACGACGACCCGAGATCAAGTCATG 960
QY 961 GGCAGGTGAGAGGCATCGGTCCCTGGATCGCGGGAGCGAGGCCCGAGTCCCGACAGGA 1020
Db 961 GGCAGGTGAGAGGCATCGGTCCCTGGATCGCGGGAGCGAGGCCCGAGTCCCGACAGGA 1020
QY 1021 GAAGACAAGACACCGGTACAGATCTGATTTATGACCGAGGGAGAAACCAATCTCGCCACC 1080
Db 1021 GAAGACAAGACACCGGTACAGATCTGATTTATGACCGAGGGAGAAACCAATCTCGCCACC 1080
QY 1081 GCAGCTACGAACCGAGCAGAGCGGAGACGAGAACCGGAGAGACACAGGCATCGAGACACCGAA 1140
Db 1081 GCAGCTACGAACCGAGCAGAGCGGAGACGAGAACCGGAGAGACACAGGCATCGAGACACCGAA 1140
QY 1141 GATCACCATCTCTGGAAGGTCTTACAAAAGAGTATAGAGATCTGGAAGGAGTTACG 1200
Db 1141 GATCACCATCTCTGGAAGGTCTTACAAAAGAGTATAGAGATCTGGAAGGAGTTACG 1200
QY 1201 GTTTATCGGTGTTCTCGAACCTGCTGGATGCAACACAGAAATTA 1260
Db 1201 GTTTATCGGTGTTCTCGAACCTGCTGGATGCAACACAGAAATTA 1260
QY 1261 AATAACAGATTTCTGGGCCCAACCCCTGGAGATTTGTGAATCATCGCTCCCAAGTAGGG 1320
Db 1261 AATAACAGATTTCTGGGCCCAACCCCTGGAGATTTGTGAATCATCGCTCCCAAGTAGGG 1320
QY 1321 AGAAGAGAGAGCTCTGTTGGAGAGAAAGAACCGTTGGAGTGAACACACAGAGTTCTG 1380
Db 1321 AGAAGAGAGAGCTCTGTTGGAGAGAAAGAACCGTTGGAGTGAACACACAGAGTTCTG 1380
QY 1381 GCAAGACAAGAACTATACCTCAATCAAGSAAAGAGCCGAGGAGACCATGCTGACA 1440
Db 1381 GCAAGACAAGAACTATACCTCAATCAAGSAAAGAGCCGAGGAGACCATGCTGACA 1440
QY 1441 AGAATGAGAGAGAGAAAGAACTTTAAAGCTGTGTGGATTCGATGCACTCATTCAG 1500
Db 1441 AGAATGAGAGAGAGAAAGAACTTTAAAGCTGTGTGGATTCGATGCACTCATTCAG 1500
QY 1501 AAACTACTACTCCAGTGACCCCATGATCAGGTGGAGATTTCTACAGTGTGGAACGA 1560
Db 1501 AAACTACTACTCCAGTGACCCCATGATCAGGTGGAGATTTCTACAGTGTGGAACGA 1560
QY 1561 GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG 1620
Db 1561 GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG 1620
QY 1621 CAAAAGCTGCTCGGCTCCGTGGGAACCTCCAAAGAAGAGCTCGATGAAGATTAGAGA 1680
Db 1621 CAAAAGCTGCTCGGCTCCGTGGGAACCTCCAAAGAAGAGCTCGATGAAGATTAGAGA 1680
QY 1681 GTTCCAGTGAAATCCGAGTGTGAGTCTGATGAGGACGACCTGTTCTAGCGCTCAGACT 1740
Db 1681 GTTCCAGTGAAATCCGAGTGTGAGTCTGATGAGGACGACCTGTTCTAGCGCTCAGACT 1740
QY 1741 CTGAAGTTTTTGAGCTTTATGAGAAATCAAAAGGAGGCGCCCTGACCGGACTTC 1800
Db 1741 CTGAAGTTTTTGAGCTTTATGAGAAATCAAAAGGAGGCGCCCTGACCGGACTTC 1800
QY 1801 ATGATGAACCTTTGGTACACGATCCAGGCCAGATGAATGAGACCACTCTGCAAAATGCA 1860
Db 1801 ATGATGAACCTTTGGTACACGATCCAGGCCAGATGAATGAGACCACTCTGCAAAATGCA 1860
QY 1861 GCGCAAGGCAAGACCGCAGGAAATTTAGGCAAGCAATTTATCTTGAGAGAGGCCCATCA 1920
|||||

Db 1861 GCGCAAGGCAAGACCGCAGGAAATTTAGGCAAGCAATTTATCTCTGGAGAGGCCCATCA 1920
QY 1921 AGCCCTGTGCTCTATGACAAACAATGCTGGGAGAGCTTTTCCACTACCGGATCACAGTCT 1980
Db 1921 AGCCCTGTGCTCTATGACAAACAATGCTGGGAGAGCTTTTCCACTACCGGATCACAGTCT 1980
QY 1981 CCGCGCTACGAACTTTTAACTGACAGAGCCCAACTGTTATAGAAATACGATGATCAGAGT 2040
Db 1981 CCGCGCTACGAACTTTTAACTGACAGAGCCCAACTGTTATAGAAATACGATGATCAGAGT 2040
QY 2041 ATATCTTTGAAGGATTTCTATGTTTGCACATGCCCCCTCGACCAATATTTCCACTGTGTA 2100
Db 2041 ATATCTTTGAAGGATTTCTATGTTTGCACATGCCCCCTCGACCAATATTTCCACTGTGTA 2100
QY 2101 AGTAATATGATTTCAACATAGACTACACGATTTCTATTTGAAGAGATGATGCGGAGA 2160
Db 2101 AGTAATATGATTTCAACATAGACTACACGATTTCTATTTGAAGAGATGATGCGGAGA 2160
QY 2161 ATTTTGTGTGAAGGCTTGAACCTTTTCTACTGTTCTTATTCAGAGATATTTTGAAT 2220
Db 2161 ATTTTGTGTGAAGGCTTGAACCTTTTCTACTGTTCTTATTCAGAGATATTTTGAAT 2220
QY 2221 TATATGACTGGAATCTTAAAGGTCTTTGTTGAAGACAGCCCTCCCTGCTGCCAAGAT 2280
Db 2221 TATATGACTGGAATCTTAAAGGTCTTTGTTGAAGACAGCCCTCCCTGCTGCCAAGAT 2280
QY 2281 TTCAATTTCAATGCGATTTTGAAGATTTCTCCAGATGGAGAAAGGATGCTGCTCCA 2340
Db 2281 TTCAATTTCAATGCGATTTTGAAGATTTCTCCAGATGGAGAAAGGATGCTGCTCCA 2340
QY 2341 TGCACAGATTTCTCTGTACTTTGTTAAAGGTGAGCAAGCCCTGGTGCCTAGGAGGAGA 2400
Db 2341 TGCACAGATTTCTCTGTACTTTGTTAAAGGTGAGCAAGCCCTGGTGCCTAGGAGGAGA 2400
QY 2401 TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGAGAGAAATGCAAG 2460
Db 2401 TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGAGAGAAATGCAAG 2460
QY 2461 GCATGATTTTACCAACCTCGGAGCAAACTGTCGTATCGATCAACTGGATC 2520
Db 2461 GCATGATTTTACCAACCTCGGAGCAAACTGTCGTATCGATCAACTGGATC 2520
QY 2521 GTGACAGTTTCAACCCCGATGTTACTTTTCCGATTTCTGCTCACTTTGGGATACGCC 2580
Db 2521 GTGACAGTTTCAACCCCGATGTTACTTTTCCGATTTCTGCTCACTTTGGGATACGCC 2580
QY 2581 CTGCAAGTTGAGTTATGACAGGAGCCCACTGACCAAACTGTGGAAGAGTTATGTGA 2640
Db 2581 CTGCAAGTTGAGTTATGACAGGAGCCCACTGACCAAACTGTGGAAGAGTTATGTGA 2640
QY 2641 AACTTGGCCACTCTCTAGCAAAATAGTCCCAAACTGACCAAACTGACCAAACTGAC 2700
Db 2641 AACTTGGCCACTCTCTAGCAAAATAGTCCCAAACTGACCAAACTGACCAAACTGAC 2700
QY 2701 CACAGAGGAGAGGCCCTCCAAAATACGCGAGAGAAATACAAATGACAGCAGAGATAA 2760
Db 2701 CACAGAGGAGAGGCCCTCCAAAATACGCGAGAGAAATACAAATGACAGCAGAGATAA 2760
QY 2761 CCGTGGAGCTAAGTAGCAGCAAGGATTTCTGGAATACTGGCAATCGTTCGTGATGTCAGC 2820
Db 2761 CCGTGGAGCTAAGTAGCAGCAAGGATTTCTGGAATACTGGCAATCGTTCGTGATGTCAGC 2820
QY 2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCGCTACCACTGCTTAATGCAAT 2880
Db 2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCGCTACCACTGCTTAATGCAAT 2880
QY 2881 TGGCAAGTTGATAGGATATCTTTTCCAAAGATGTTGTTGCTGAGCTGGGCACTGACTC 2940
Db 2881 TGGCAAGTTGATAGGATATCTTTTCCAAAGATGTTGTTGCTGAGCTGGGCACTGACTC 2940
QY 2941 ATCCAGATCATCATTTTAAATTTTGAATGAAATCCCTGATCATGCGAGAAATTCATTTCTA 3000
Db 2941 ATCCAGATCATCATTTTAAATTTTGAATGAAATCCCTGATCATGCGAGAAATTCATTTCTA 3000

QY 3001 ACTGTGGAATTCGGCAGCCCAAATACGGAGACAGAAAAGTTTCATCACATGCACATGCGGA 3060
Db
QY 3001 ACTGTGGAATTCGGCAGCCCAAATACGGAGACAGAAAAGTTTCATCACATGCACATGCGGA 3060
Db
QY 3061 AGAAGGGGATTAACACCTTTGATAAATATCATGTGACGCCCTTGCCCAAGATGACCCCAATC 3120
Db
QY 3061 AGAAGGGGATTAACACCTTTGATAAATATCATGTGACGCCCTTGCCCAAGATGACCCCAATC 3120
QY 3121 CCTCGAGGATTAACCACAATGAAACGTTTGGAAATTCCTGGGTGATGCTGTTGTAATTC 3180
Db
QY 3121 CCTCGAGGATTAACCACAATGAAACGTTTGGAAATTCCTGGGTGATGCTGTTGTAATTC 3180
QY 3181 TGACCAAGCGTTCATTTGTAATAATTCCTGCTGCTGAGTCTGGAAGAGGAGGATTAGCAACT 3240
Db
QY 3181 TGACCAAGCGTTCATTTGTAATAATTCCTGCTGCTGAGTCTGGAAGAGGAGGATTAGCAACT 3240
QY 3241 ATCGGACTGCCATTTGTCAGAATCAGACCTTTGCCATGCTAGCAAGAAGAACTTGAACCTGG 3300
Db
QY 3241 ATCGGACTGCCATTTGTCAGAATCAGACCTTTGCCATGCTAGCAAGAAGAACTTGAACCTGG 3300
QY 3301 ATCCATTTATGCTGATGCTCAGCGGCCCTGACCTTTGTGTAGAGAAATCGGACCTTCGACATG 3360
Db
QY 3301 ATCCATTTATGCTGATGCTCAGCGGCCCTGACCTTTGTGTAGAGAAATCGGACCTTCGACATG 3360
QY 3361 CAATGGCCAAATGTTTTGAAGCGTTAATAGGAGCTGTTTACTTGGAGGGAAGCCTGGAGG 3420
Db
QY 3361 CAATGGCCAAATGTTTTGAAGCGTTAATAGGAGCTGTTTACTTGGAGGGAAGCCTGGAGG 3420
QY 3421 AAGCCAAAGCAGTTATTTGGACGCTTCTCTTTAAATGATCCGACCTTCGCGAAGTCTGGC 3480
Db
QY 3421 AAGCCAAAGCAGTTATTTGGACGCTTCTCTTTAAATGATCCGACCTTCGCGAAGTCTGGC 3480
QY 3481 TCAATTTATCCTCTCCACCCCACTCCAACTACAGAGCCAAATACGTGATCGACAACTTATTTG 3540
Db
QY 3481 TCAATTTATCCTCTCCACCCCACTCCAACTACAGAGCCAAATACGTGATCGACAACTTATTTG 3540
QY 3541 AAATTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATGGAGTAATTTTGA 3600
Db
QY 3541 AAATTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATGGAGTAATTTTGA 3600
QY 3601 CTATGTTTCGACTCTGCGCAAGGGCAATTCACATTTAGAACTGTGGGATTTAAACCACTGA 3660
Db
QY 3601 CTATGTTTCGACTCTGCGCAAGGGCAATTCACATTTAGAACTGTGGGATTTAAACCACTGA 3660
QY 3661 CCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAATCGGTAGCCA 3720
Db
QY 3661 CCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAATCGGTAGCCA 3720
QY 3721 CAGAGTACTTATTTATTTATTTCCAGATCATCATGAAGGACACTTAACTTTGTTGCGAA 3780
Db
QY 3721 CAGAGTACTTATTTATTTATTTCCAGATCATCATGAAGGACACTTAACTTTGTTGCGAA 3780
QY 3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGAGCTGGGCACTGAGAGT 3840
Db
QY 3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGAGCTGGGCACTGAGAGT 3840
QY 3841 AGCCCAATACCAACGACAAAGACCGCTGTGCGCTTCGCAACAAGACCTTGGCGG 3900
Db
QY 3841 AGCCCAATACCAACGACAAAGACCGCTGTGCGCTTCGCAACAAGACCTTGGCGG 3900
QY 3901 ACCTTTGGAAATCATTTATTTGACCGCTGTACACTGATAGGATTTGGAAATATGTTTATA 3960
Db
QY 3901 ACCTTTGGAAATCATTTATTTGACCGCTGTACACTGATAGGATTTGGAAATATGTTTATA 3960
QY 3961 CTTTCATGAATGCTGCTCTTTCCAGATTTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
Db
QY 3961 CTTTCATGAATGCTGCTCTTTCCAGATTTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
QY 4021 ATGACCCCAAAATCCCAAGCTTCAGACGTGTTGTCACCTTAGGACAGAGGAAAGAGC 4080
Db
QY 4021 ATGACCCCAAAATCCCAAGCTTCAGACGTGTTGTCACCTTAGGACAGAGGAAAGAGC 4080

QY 4081 CAGACATTTCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTACA 4140
Db
QY 4081 CAGACATTTCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTACA 4140
QY 4141 CTGTGGCTGTTTATTTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Db
QY 4141 CTGTGGCTGTTTATTTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
QY 4201 AAGCCGAAATGGAGCAGCAATGATGCTTGAATAATATATATTTTCCCAGATGGCCC 4260
Db
QY 4201 AAGCCGAAATGGAGCAGCAATGATGCTTGAATAATATATATTTTCCCAGATGGCCC 4260
QY 4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGAACAGAGTTTAAAGAAATGAGGTGGAAA 4320
Db
QY 4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGAACAGAGTTTAAAGAAATGAGGTGGAAA 4320
QY 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGGAAGACATCAAGAAATAAAGAGGGCAT 4380
Db
QY 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGGAAGACATCAAGAAATAAAGAGGGCAT 4380
QY 4381 GCAAGTGTGGAGTATTTTACTTGTCTCAGTAACTGTGCACTGTTCTTATTGAGACCTAGCCT 4440
Db
QY 4381 GCAAGTGTGGAGTATTTTACTTGTCTCAGTAACTGTGCACTGTTCTTATTGAGACCTAGCCT 4440
QY 4441 AGTTTTTCTCGCAGACAAATGAAAGTGTGCTCATTGAAATAAAATACAGAGTCAAATCG 4500
Db
QY 4441 AGTTTTTCTCGCAGACAAATGAAAGTGTGCTCATTGAAATAAAATACAGAGTCAAATCG 4500
QY 4501 CTATTTGTTTAAATGATCTGTTTTTGTAGTGGATGGTCTTTATTACAAAGTATTAGATT 4560
Db
QY 4501 CTATTTGTTTAAATGATCTGTTTTTGTAGTGGATGGTCTTTATTACAAAGTATTAGATT 4560
QY 4561 TTTCTCTTATTTAAACGGAACCTTCGACTTTGGTGAATGTGCATTTCTCTTTTATTG 4620
Db
QY 4561 TTTCTCTTATTTAAACGGAACCTTCGACTTTGGTGAATGTGCATTTCTCTTTTATTG 4620
QY 4621 CTCCTTAAATAATAAAATTCAGAAAGCATATTCTATGTGGAATAGATCCTGTTTTCCAT 4680
Db
QY 4621 CTCCTTAAATAATAAAATTCAGAAAGCATATTCTATGTGGAATAGATCCTGTTTTCCAT 4680
QY 4681 CTGTGCTCCAGATTTGTGACCTTAGACTTTTCAATTTGACAAAGTAAAAATTGACTTTACTAG 4740
Db
QY 4681 CTGTGCTCCAGATTTGTGACCTTAGACTTTTCAATTTGACAAAGTAAAAATTGACTTTACTAG 4740
QY 4741 TAAAAAATAAAAAAAAAAAAAA 4764
Db
QY 4741 TAAAAAATAAAAAAAAAAAAAA 4764

RESULT 6
ADR68726
ID ADR68726 standard; cDNA; 4764 BP.
XX
AC ADR68726;
XX AC
DT 02-DEC-2004 (first entry)
XX
DE Human cDNA encoding Ribonuclease III, RNase III.
XX
KW Human; ss; gene; Ribonuclease III; RNase III; RNA interference;
KW gene silencing; double stranded RNA; pre-rRNA processing; RNA processing;
KW RNA expression; RNA splicing; RNA translocation.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 246..4370
FT /*tag= a
FT /product= "Ribonuclease III"
FT /transl_except= (pos:3876..3878,aa:Gly)
FT /transl_except= (pos:4278..4280,aa:Gly)
XX
PN US2004175828-A1.

XX 09-SEP-2004.
XX PD
XX 22-MAR-2004; 2004US-00805919.
XX PF
XX 06-JUL-2001; 2001US-00900425.
XX PR
XX (CROO/) CROOKE S T.
XX PA (WUHH/) WU H.
XX XX
XX Crooke ST, Wu H;
XX PI
XX WPI; 2004-667111/65.
XX DR
XX Modulating RNA interference, processing, expression, splicing and
XX PT translocation in cell or tissue, by contacting cell or tissue with human
XX PT RNase III or oligomeric compound targeted to nucleic acid encoding human
XX PT RNase III.
XX PT
XX PS Example 2; SEQ ID NO 1; 32pp; English.
XX CC The invention relates to modulating RNA interference (gene silencing),
XX CC RNA processing (e.g. of pre-rRNA), RNA expression, RNA splicing, or RNA
XX CC translocation in a cell or tissue, by contacting the cell or tissue with
XX CC a modulator effective to cause the modulation by at least 50% as compared
XX CC to control, where modulator is human RNase III (ribonuclease III, which
XX CC cleaves double stranded RNA) polypeptide or an oligomeric compound (an
XX CC antisense oligonucleotide) targeted to the nucleic acid encoding human
XX CC RNase III. The oligomeric compound is targeted to a 3'-untranslated
XX CC region (3'UTR), a 5'-untranslated region (5'UTR) or a coding region of a
XX CC nucleic acid molecule encoding human RNase III polypeptide, where the
XX CC oligomeric compound inhibits the expression of human RNase III
XX CC polypeptide by at least 50%. In the method above, the RNA is rRNA, snRNA,
XX CC snoRNA or miRNA, or precursors of the above. The RNA is processed to form
XX CC one or more 30S and 32S RNA fragments. The 32S RNA is processed to form
XX CC one or more 12S pre-rRNA and 28S rRNA fragments. The methods are useful
XX CC for modulating RNA interference in a cell or tissue, modulating
XX CC processing of an RNA in a cell or tissue, modulating RNA expression in a
XX CC cell or tissue, modulating RNA splicing in a cell or tissue, and
XX CC modulating RNA translocation in a cell or tissue. The methods are useful
XX CC for research, biological and clinical purposes. The methods are useful in
XX CC defining the roles of RNase III and interaction of human RNase III and
XX CC cellular RNA. The present sequence encodes human RNaseIII.
XX SQ Sequence 4764 BP; 1379 A; 1138 C; 1082 G; 1165 T; 0 U; 0 Other;
Query Match 99.9%; Score 4760.8; DB 13; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CTGCTTGTAGCTCGCGTAGTAGCCCTGGCTTGTCTCTGACGCGCATCTCGGCGCCGAG 60
DB 1 CTGCTTGTAGCTCGCGTAGTAGCCCTGGCTTGTCTCTGACGCGCATCTCGGCGCCGAG 60
QY 61 AGCCTTTTATAGTTCTTTCGCGGAGTGTGAGGATACAGAAATGATGTGAATCAA 120
DB 61 AGCCTTTTATAGTTCTTTCGCGGAGTGTGAGGATACAGAAATGATGTGAATCAA 120
QY 121 CCATATCATCAGGAGCTGTAATCTAGTGAAGAGTTAGAGTGTGCATCTTCACTA 180
DB 121 CCATATCATCAGGAGCTGTAATCTAGTGAAGAGTTAGAGTGTGCATCTTCACTA 180
QY 181 TGATATGAGCGCAGTCTCTGAGCTTATATTTCTGTGTGAAGATGTGACATATCCAGCGGA 240
DB 181 TGATATGAGCGCAGTCTCTGAGCTTATATTTCTGTGTGAAGATGTGACATATCCAGCGGA 240
QY 241 ACATCATGATGAGGAAACACATGTCA CAGAAATGTCTTCCACCGGAGCAGGCGCTC 300
DB 241 ACATCATGATGAGGAAACACATGTCA CAGAAATGTCTTCCACCGGAGCAGGCGCTC 300
QY 301 CCGAGGACGAGGAGACATGGACCGACACCTCTCAGCACCATCTTTAGGCCCCCAATC 360
DB 301 CCGAGGACGAGGAGACATGGAGCCGACACCTCTCAGCACCATCTTTAGGCCCCCAATC 360

QY 361 TGAGGCTGCTTCA CCGCTCAGCAGCCTCTGTGCAATATCAATATGAACTTCAAGTGCC 420
DB 361 TGAGGCTGCTTCA CCGCTCAGCAGCCTCTGTGCAATATCAATATGAACTTCAAGTGCC 420
QY 421 CTTCCACCACTTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACACAGACTTTG 480
DB 421 CTTCCACCACTTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACAGACTTTG 480
QY 481 TACCTTTCCCCCACCACCATGCTCCGTGAGCGAAGGCCCTCTTCCCCCTGCCCAATCA 540
DB 481 TACCTTTCCCCCACCACCATGCTCCGTGAGCGAAGGCCCTCTTCCCCCTGCCCAATCA 540
QY 541 GGCCTCTTTCCCAACACAGATGAGGCA CCGCTTCCAGTTCTCTCTTTCTTTTCTCCTC 600
DB 541 GGCCTCTTTCCCAACACAGATGAGGCA CCGCTTCCAGTTCTCTCTTTCTTTTCTCCTC 600
QY 601 CCATGCGCACCAATGCTGTCTTAATAA CCCCCCAGTCCCTGGGGCACCTCTCTGGAC 660
DB 601 CCATGCGCACCAATGCTGTCTTAATAA CCCCCCAGTCCCTGGGGCACCTCTCTGGAC 660
QY 661 AAGGCACTTTTCCCTTCATGATGCCCT CCGCTCTCATGCTCATCCCCCGCCCTCCAG 720
DB 661 AAGGCACTTTTCCCTTCATGATGCCCT CCGCTCTCATGCTCATCCCCCGCCCTCCAG 720
QY 721 TCATGCGCAGCAGGTTAATTAATCATG TACCTCCCGGCTATTCTACCACTTCCAC 780
DB 721 TCATGCGCAGCAGGTTAATTAATCATG TACCTCCCGGCTATTCTACCACTTCCAC 780
QY 781 CTCCTCAGTTTAAATAGTTTTCAGAA CACCCTAGTTCTTTCTGCGCCAGTGTCTAATA 840
DB 781 CTCCTCAGTTTAAATAGTTTTCAGAA CACCCTAGTTCTTTCTGCGCCAGTGTCTAATA 840
QY 841 GCAGTAGTCTCTCATTTTCAGACATCT CCGCTCCATACCCACTCCCAAGGCTCCAGTGAG 900
DB 841 GCAGTAGTCTCTCATTTTCAGACATCT CCGCTCCATACCCACTCCCAAGGCTCCAGTGAG 900
QY 901 GAAGTCCCCAGAAAGGCTGAAACATATG ATGACACAGGACCGAGACACAGTCATG 960
DB 901 GAAGTCCCCAGAAAGGCTGAAACATATG ATGACACAGGACCGAGATCAAGTCATG 960
QY 961 GCGGAGGTGAGAGGATCGTCCCTTGGAT CCGCGGAGCGGCGCGCTCCGACAGGA 1020
DB 961 GCGGAGGTGAGAGGATCGTCCCTTGGAT CCGCGGAGCGGCGCGCTCCGACAGGA 1020
QY 1021 GAAAGACAAGACAGCGGTACAGATCTGA TTTATGACCGAGGAGAACACCATCTCGCCACC 1080
DB 1021 GAAAGACAAGACAGCGGTACAGATCTGA TTTATGACCGAGGAGAACACCATCTCGCCACC 1080
QY 1081 GCAGCTACGAAACGAGACGAGACGAGAA CCGGAGAGACACAGGATCGAGACAACCGAA 1140
DB 1081 GCAGCTACGAAACGAGACGAGACGAGAA CCGGAGAGACACAGGATCGAGACAACCGAA 1140
QY 1141 GATCACCATCTCTGGAAGGTCCTACAAA AAGAGTATAAGAGATCTGGAAGAGTTACG 1200
DB 1141 GATCACCATCTCTGGAAGGTCCTACAAA AAGAGTATAAGAGATCTGGAAGAGTTACG 1200
QY 1201 GTTTATCGGTTTCTCTGAACTCTGAA CCGTGTGATGCACACAGAAATTACCTGGGAGATTATTA 1260
DB 1201 GTTTATCGGTTTCTCTGAACTCTGAA CCGTGTGATGCACACAGAAATTACCTGGGAGATTATTA 1260
QY 1261 AAAATACAGATTCTTTGGGCCCCCACC CCGTGTGAGATTGTGAATCATCTGCTCCCAAGTAGGG 1320
DB 1261 AAAATACAGATTCTTTGGGCCCCCACC CCGTGTGAGATTGTGAATCATCTGCTCCCAAGTAGGG 1320
QY 1321 AGAAGACAGAGCTGTTGGGAGGAGAAA AAGACCGTTGGAGTGACACACAGAGTTCTG 1380
DB 1321 AGAAGACAGAGCTGTTGGGAGGAGAAA AAGACCGTTGGAGTGACACACAGAGTTCTG 1380
QY 1381 GCAAGACAGAAATATATACCTCAATCA AAGAAAAAGAGCCCGAGGAGACCATCTGCTGACA 1440
DB 1381 GCAAGACAGAAATATATACCTCAATCA AAGAAAAAGAGCCCGAGGAGACCATCTGCTGACA 1440

1441 AGAATGAGGAGGAAGAAGAACTTTCTTAAGCCTGTGTGGATTTCGATGCACCTCATTTGAG 1500
1441 AGAATGAGGAGGAAGAAGAACTTTCTTAAGCCTGTGTGGATTTCGATGCACCTCATTTGAG 1500
1501 AAAAATCTATCTCCAGTGAGCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAAAGCA 1560
1501 AAAAATCTATCTCCAGTGAGCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAAAGCA 1560
1561 GTAGGCTTGTGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAAGG 1620
1561 GTAGGCTTGTGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAAGG 1620
1621 CCAAGCTGCTCGGCTCCTCGTGGGAACCTCCAAAGACGAAGCTTCGATGAAGATTTAGAGA 1680
1621 CCAAGCTGCTCGGCTCCTCGTGGGAACCTCCAAAGACGAAGCTTCGATGAAGATTTAGAGA 1680
1681 GTTCCAGTGAATCCGAGTGTGAGTGTGATGAGGACAGCACTCTGTTCAGAGCTCAGACT 1740
1681 GTTCCAGTGAATCCGAGTGTGAGTGTGATGAGGACAGCACTCTGTTCAGAGCTCAGACT 1740
1741 CTGAAGCTTTTGACGTTTATGCAAGAAATCAAAAGCGCAAAAGGCCCAACCTGACCGACTTC 1800
1741 CTGAAGCTTTTGACGTTTATGCAAGAAATCAAAAGCGCAAAAGGCCCAACCTGACCGACTTC 1800
1801 ATGATGAACCTTTGGTACAAAGATCCAGGCGCAGATGAATGATGACCACTCTGCAAAATGCA 1860
1801 ATGATGAACCTTTGGTACAAAGATCCAGGCGCAGATGAATGATGACCACTCTGCAAAATGCA 1860
1861 GGCAAAAGGCAAGACGCAAGGAATTTAGGCAAGCAATTTATCTGAGGAAGAGGCCCATCA 1920
1861 GGCAAAAGGCAAGACGCAAGGAATTTAGGCAAGCAATTTATCTGAGGAAGAGGCCCATCA 1920
1921 AGCCCTGTGCTCTATGACCAAAATGCTGGGAGACTTTTCCACTACCGGATCAGACTCT 1980
1921 AGCCCTGTGCTCTATGACCAAAATGCTGGGAGACTTTTCCACTACCGGATCAGACTCT 1980
1981 CCCCCTGTGAAAGGCTTTGAACTCTTTTACGACGCGCACTGTTATAGAAATACGATGATCAGAGT 2040
1981 CCCCCTGTGAAAGGCTTTGAACTCTTTTACGACGCGCACTGTTATAGAAATACGATGATCAGAGT 2040
2041 ATATCTTTGAAAGGATTTCTATGTTTGCACATGCCCCCTGACCAATATTTCCAATGTGTA 2100
2041 ATATCTTTGAAAGGATTTCTATGTTTGCACATGCCCCCTGACCAATATTTCCAATGTGTA 2100
2101 AAGTAATAGATTTCAACATAGACTACAGATTCATTTCAATTTGAAGAGATGATGCGCGAGA 2160
2101 AAGTAATAGATTTCAACATAGACTACAGATTCATTTCAATTTGAAGAGATGATGCGCGAGA 2160
2161 ATTTTGTGTGAAAGGCTTTGAACTCTTTTACGACGCGCACTGTTTACGAGATATTTTGGAAAT 2220
2161 ATTTTGTGTGAAAGGCTTTGAACTCTTTTACGACGCGCACTGTTTACGAGATATTTTGGAAAT 2220
2221 TATATGACTGGAACTCTTAAAGGTCCTTTGTTGAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
2221 TATATGACTGGAACTCTTAAAGGTCCTTTGTTGAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
2281 TTCAATTTATGCGACGCTTTTGAAGATTTCTCCAGATGAGGAAAGGAAGTCTGTCCA 2340
2281 TTCAATTTATGCGACGCTTTTGAAGATTTCTCCAGATGAGGAAAGGAAGTCTGTCCA 2340
2341 TGCACAGATTTCTCTGTACTCTTAAAGGTGAGCAAGCCCTGGTGCCTGAGGAGGAGA 2400
2341 TGCACAGATTTCTCTGTACTCTTAAAGGTGAGCAAGCCCTGGTGCCTGAGGAGGAGA 2400
2401 TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGAGGAAGATGCAAAAG 2460
2401 TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGAGGAAGATGCAAAAG 2460
2461 GCATGATTTGTACCAACCCCTGGGACGAACCAAGCTCTGTCCGATTCGATCAACTGGATC 2520
2461 GCATGATTTGTACCAACCCCTGGGACGAACCAAGCTCTGTCCGATTCGATCAACTGGATC 2520
2521 GTGAACAGTTTCAACCCCGATGTGATTTACTTTTCCGATTTATCGTCCACTTTTGGGATACGCC 2580

2521 GTGAACAGTTTCAACCCCGATGTGATTTACTTTTCCGATTTATCGTCCACTTTTGGATACGCC 2580
2581 CTGCAAGTTGAGTTTATGCAAGGAGACCCACAGTACCAAAAACTGTGGAAGAGTTATGTGA 2640
2581 CTGCAAGTTGAGTTTATGCAAGGAGACCCACAGTACCAAAAACTGTGGAAGAGTTATGTGA 2640
2641 AACTTTCGCACCTCCTAGCAAAATAGTCCCAAAAGTCAAAACAACTGCAAAACAGAAAGCTGG 2700
2641 AACTTTCGCACCTCCTAGCAAAATAGTCCCAAAAGTCAAAACAACTGCAAAACAGAAAGCTGG 2700
2701 CACAGAGGAGAAAGCCCTCCAAAAATAACGGCAGAGAAATACAAATGAGACAGAAAGTAA 2760
2701 CACAGAGGAGAAAGCCCTCCAAAAATAACGGCAGAGAAATACAAATGAGACAGAAAGTAA 2760
2761 CGGTGGAGTAAGTAGCCAAAGGATTTCTGGAACCTGGCAATCCGTTCTGATGCTGTGCAGC 2820
2761 CGGTGGAGTAAGTAGCCAAAGGATTTCTGGAACCTGGCAATCCGTTCTGATGCTGTGCAGC 2820
2821 ATGCAATGATGCTACTGTTCTGACCCATCATATCCGCTACCAACCAATGCCTAAATGCATT 2880
2821 ATGCAATGATGCTACTGTTCTGACCCATCATATCCGCTACCAACCAATGCCTAAATGCATT 2880
2881 TGGACAAGTTGATAGATATATCTTCCAAAGATGCTGTCTGTTCGACAGCTGGCCATGACTC 2940
2881 TGGACAAGTTGATAGATATATCTTCCAAAGATGCTGTCTGTTCGACAGCTGGCCATGACTC 2940
2941 ATCCAAGTCATCATTTAAATTTTGGAAATGAAATCCATGATCATGCCAGGAATTCATTAATCTA 3000
2941 ATCCAAGTCATCATTTAAATTTTGGAAATGAAATCCATGATCATGCCAGGAATTCATTAATCTA 3000
3001 ACTGTGGAATTCGGCAGCCCAAAATACGGAGACAGAAAAAGTTTCATCATGTCACATCGGGA 3060
3001 ACTGTGGAATTCGGCAGCCCAAAATACGGAGACAGAAAAAGTTTCATCATGTCACATCGGGA 3060
3061 AGAAAAGGATTAACACCTTTGATPAAATATCATGTACGCTTGGCCCAAGATGACCCAACTC 3120
3061 AGAAAAGGATTAACACCTTTGATPAAATATCATGTACGCTTGGCCCAAGATGACCCAACTC 3120
3121 CCTCGAGGATTAACCAATGAAACGGTTGGAATTCCTGGGTGATGCTGTTGTGAATTTTC 3180
3121 CCTCGAGGATTAACCAATGAAACGGTTGGAATTCCTGGGTGATGCTGTTGTGAATTTTC 3180
3181 TGACACAGCTCCATTTGTACTATTTGTTTCTAGTCTGCAAGAAAGAGAGATTAGCAACCT 3240
3181 TGACACAGCTCCATTTGTACTATTTGTTTCTAGTCTGCAAGAAAGAGAGATTAGCAACCT 3240
3241 ATCGGACTGCCATTTGTTAGAAATCAGACCTTTGCCATGCTAGCAAAAGAACTTTGAACTGG 3300
3241 ATCGGACTGCCATTTGTTAGAAATCAGACCTTTGCCATGCTAGCAAAAGAACTTTGAACTGG 3300
3301 ATCCATTTATGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACCTTTGACATG 3360
3301 ATCCATTTATGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACCTTTGACATG 3360
3361 CAATGCCCCTTTGTTTGAAGGCTTAAATAGGAGCTGTTTACTTTGAGGGAAGCTCGAGG 3420
3361 CAATGCCCCTTTGTTTGAAGGCTTAAATAGGAGCTGTTTACTTTGAGGGAAGCTCGAGG 3420
3421 AAGCCAAAGCAGTTATTTGGACGCTTCTCTTAAATGATCCGACCTTCGCGAAGTCTGGC 3480
3421 AAGCCAAAGCAGTTATTTGGACGCTTCTCTTAAATGATCCGACCTTCGCGAAGTCTGGC 3480
3481 TCATTTATCTCTCCACCCACTCCAACTACAGAGCCAAATCTGATCCGACCTTTATTG 3540
3481 TCATTTATCTCTCCACCCACTCCAACTACAGAGCCAAATCTGATCCGACCTTTATTG 3540
3541 AAACCTTCTCCAGTTCTACAAAAAATTTACTGAGTTTGAAGAAGCAATTTGGAGTAATTTTA 3600
3541 AAACCTTCTCCAGTTCTACAAAAAATTTACTGAGTTTGAAGAAGCAATTTGGAGTAATTTTA 3600
3601 CTCAGTTTCGACTTCTGGCAAGGGCAATTCATTTGAGAACTGTGGGATTTAACCATCTGA 3660
3601 CTCAGTTTCGACTTCTGGCAAGGGCAATTCATTTGAGAACTGTGGGATTTAACCATCTGA 3660

Db 3601 CTGATGTTGACTTCTGGCAAGGCATTTCATTGAGAACTGTGGGATTTAACCATCTGA 3660
Qy 3661 CCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATCAACTGCTGTAGCCA 3720
Db 3661 CCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATCAACTGCTGTAGCCA 3720
Qy 3721 CAGAGTACTTATTCATCTTCCAGATCATCATGAGGACACTTAACTTTGTCGGA 3780
Db 3721 CAGAGTACTTATTCATCTTCCAGATCATCATGAGGACACTTAACTTTGTCGGA 3780
Qy 3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGCGATGCGAGGT 3840
Db 3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGCGATGCGAGGT 3840
Qy 3841 ACGCCATAACAACGACAGCAAGAGGCTGTGGCGCTTCGCAACAGACCTTGGCGG 3900
Db 3841 ACGCCATAACAACGACAGCAAGAGGCTGTGGCGCTTCGCAACAGACCTTGGCGG 3900
Qy 3901 ACCTTTTGGATCAATTTATTCAGCGCTGTACATGATAGGATTTGGAATATGTTTCA 3960
Db 3901 ACCTTTTGGATCAATTTATTCAGCGCTGTACATGATAGGATTTGGAATATGTTTCA 3960
Qy 3961 CTTTCATGAATGTCTGCTTTCCAGATTTGAAGAAATTCATTTGAATCAGATTTGGA 4020
Db 3961 CTTTCATGAATGTCTGCTTTCCAGATTTGAAGAAATTCATTTGAATCAGATTTGGA 4020
Qy 4021 ATGACCCCAATCCAGCTTCAGCAGTGTGCTTGACATTTAGGACAGAGAAAGAGC 4080
Db 4021 ATGACCCCAATCCAGCTTCAGCAGTGTGCTTGACATTTAGGACAGAGAAAGAGC 4080
Qy 4081 CAGACATCTCTGTACAGACTCTGCAGACAGTGGGCCCCATCCCATGCCCGAACCTAGA 4140
Db 4081 CAGACATCTCTGTACAGACTCTGCAGACAGTGGGCCCCATCCCATGCCCGAACCTAGA 4140
Qy 4141 CTGTGGCTGTTTATTTCAAGGAGAAAGATAGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Db 4141 CTGTGGCTGTTTATTTCAAGGAGAAAGATAGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Qy 4201 AAGCGGAAATGGGAGCAGCAATGATGCGCTTGAAATAATATTTTCCAGATGCCCC 4260
Db 4201 AAGCGGAAATGGGAGCAGCAATGATGCGCTTGAAATAATATTTTCCAGATGCCCC 4260
Qy 4261 ATCAGAAGCGGTTTCATCGAAGGAGTACAGACAGAGTTAAAGAAATGAGTGGGAAA 4320
Db 4261 ATCAGAAGCGGTTTCATCGAAGGAGTACAGACAGAGTTAAAGAAATGAGTGGGAAA 4320
Qy 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGGAGGCGAT 4380
Db 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGGAGGCGAT 4380
Qy 4381 GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
Db 4381 GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
Qy 4441 AGTTTCTCAGACAAATGAACGAGTGTCTCATTTGAATAAATAACAGAGTCAATCG 4500
Db 4441 AGTTTCTCAGACAAATGAACGAGTGTCTCATTTGAATAAATAACAGAGTCAATCG 4500
Qy 4501 CTATTGTTGTTTAAATGATCTGTTTATGCTGATGCTCTTATTACAAAGTATTAGATT 4560
Db 4501 CTATTGTTGTTTAAATGATCTGTTTATGCTGATGCTCTTATTACAAAGTATTAGATT 4560
Qy 4561 TTTCTTCTATTAAACGGAAATCTGACTTTGGTGAATGTGCATTAATCTCTTTTATTTTG 4620
Db 4561 TTTCTTCTATTAAACGGAAATCTGACTTTGGTGAATGTGCATTAATCTCTTTTATTTTG 4620
Qy 4621 CTCTTTAAATPAATAAATCAAGAGCATATTTATGTTGAAATAGATCTGTTTTCAT 4680
Db 4621 CTCTTTAAATPAATAAATCAAGAGCATATTTATGTTGAAATAGATCTGTTTTCAT 4680
Qy 4681 CTGTGTCAGATGTGACCTAGACTTCAATTTGACAGTAAATAATGACTTTACTAG 4740
Db 4681 CTGTGTCAGATGTGACCTAGACTTCAATTTGACAGTAAATAATGACTTTACTAG 4740

Qy 4741 TAAAAA 4764
Db 4741 TAAAAA 4764

RESULT 7

ADX03768

ID ADX03768 standard; cDNA; 4764 BP.

XX ADX03768;

XX 05-MAY-2005 (first entry)

XX Human cDNA encoding Drosha RNase III.

DE Drosha; RNase III; RNA interference; gene silencing; Cytostatic;
XX Antidiabetic; Anorectic; Antilipemic; Antiarteriosclerotic; Hypotensive;
KW Neuroprotective; Nootropic; Antiangiogenic; Anabolic;
KW Eating-Disorders-Gen.; hyperproliferation; cancer; neoplasm;
KW angiogenesis disorder; cardiovascular disease;
KW non-insulin dependent diabetes; endocrine disease;
KW gastrointestinal disease; metabolic disorder; obesity;
KW nutritional disorder; hyperlipidemia; atherosclerosis; atherogenesis;
KW hypertension; anorexia nervosa; nutritional disorder;
KW psychiatric disorder; Alzheimers disease; degeneration;
KW neurological disease; nervous system injury; neurodegenerative disease;
KW neurological disorder; gene; ss.

OS Homo sapiens.

XX W02005013901-A2.

XX 17-FEB-2005.

XX 30-JUL-2004; 2004WO-US025300.

XX 31-JUL-2003; 2003US-0492056P.

PR 31-OCT-2003; 2003US-0516303P.

PR 19-DEC-2003; 2003US-0531596P.

PR 14-APR-2004; 2004US-0562417P.

XX (ISIS-) ISIS PHARM INC.

XX Esau C, Lollo B, Bennett CF, Freier SM, Griffey RH, Baker BF;

PI Vickers T, Marcussen EG, Koller E, Swayze EE, Jain R, Bhat B;

PI Peralta E;

XX WPI; 2005-163123/17.

XX Example 24; SEQ ID NO 860; 854pp; English.

PT New oligomeric compound that can hybridize with or sterically interfere
PT with nucleic acid molecules comprising or encoding small non-coding RNA
PT targets, useful for treating e.g., cancer and diabetes.
XX
XX The invention relates to an oligomeric compound comprising a first region
CC and a second region, where at least one region contains a modification,
CC and a portion of the oligomeric compound is targeted to a small non-
CC coding RNA target nucleic acid that is miRNA (micro-RNA), or its
CC precursor (primary-miRNA, pri-miRNA). Also included are a composition
CC comprising a first oligomeric compound and a second oligomeric compound
CC (where at least one of the oligomeric compounds contains a modification,
CC at least a portion of the first oligomeric compound is capable of
CC hybridizing with at least a portion of the second oligomeric compound,
CC and at least a portion of the first oligomeric compound is targeted to a
CC small non-coding RNA target nucleic acid), a pharmaceutical composition
CC comprising the composition cited above (and a carrier), a kit or assay
CC device comprising the composition, modulating the expression of a small
CC non-coding RNA target nucleic acid in a cell (or tissue or animal),
CC treating or preventing a disease or disorder associated with a small non-
CC coding RNA target nucleic acid, treating a condition in an animal,
CC treating or preventing a disease or disorder associated with CD36,

CC methods of screening an oligomeric compound for an effect on miRNA
CC signaling, methods of screening a miRNA precursor for an effect in miRNA
CC signaling, methods of modulating translation (or apoptosis, conversion of
CC a precursor miRNA into miRNA, or cellular differentiation), identifying
CC an RNA transcript bound to a small non-coding RNA, arresting (or
CC delaying) entry of a cell at the G2/M phase, interfering with chromosome
CC segregation, a method of triggering apoptosis, detecting a miRNA
CC precursor, identifying a miRNA target, modulating cellular
CC differentiation, treating a condition associated with adipocyte
CC differentiation in an animal, treating/preventing a disease/disorder
CC associated with aberrant regulation of the cell cycle by miRNAs,
CC maintaining a pluripotent stem cell and identifying a small non-coding
CC RNA binding site. The oligomeric compound is targeted to a region
CC flanking a Drosha cleavage site within a pri-miRNA. It stimulates an
CC increase in expression of a pri-miRNA. The compounds and compositions are
CC useful for treating a disease or disorder resulting from chromosomal non-
CC disjunction, altered methylation, acetylation, or pseudouridylation state
CC of chromosomes, such as a hyperproliferative condition (e.g. cancer,
CC neoplasia or angiogenesis), diabetes (Type 2 diabetes), obesity,
CC hyperlipidemia, atherosclerosis, atherogenesis, hypertension, anorexia,
CC Alzheimer's disease, a central nervous system injury or neurodegenerative
CC disorder. The present sequence is a cDNA representing the Drosha gene, an
CC RNase III that processes pri-miRNA into 70bp pre-miRNA (subsequently
CC processed to miRNA by Dicer RNase).

XX
SQ Sequence 4764 BP; 1379 A; 1138 C; 1082 G; 1165 T; 0 U; 0 Other;

Query Match 99.9%; Score 4760.8; DB 14; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGCTTGTGTACCTCGGTAGTACCTGGCTTTGCTCTGACGCGCATCTCGGCGCCGAG 60
DB |||||
QY 1 CTGCTTGTGTACCTCGGTAGTACCTGGCTTTGCTCTGACGCGCATCTCGGCGCCGAG 60
DB |||||

QY 61 AGCCTTTTATAGTGTCTTTCCGCGGATGTGAAGATACAGAAATGACTGTGAATCAA 120
DB |||||

QY 61 AGCCTTTTATAGTGTCTTTCCGCGGATGTGAAGATACAGAAATGACTGTGAATCAA 120
DB |||||

QY 121 CCATATCATCAAGGAGCTGATTAATCTAGTGAAGAGTTAGCGTGTGCATCTCACTA 180
DB |||||

QY 121 CCATATCATCAAGGAGCTGATTAATCTAGTGAAGAGTTAGCGTGTGCATCTCACTA 180
DB |||||

QY 181 TGATATGAGCGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGCGGA 240
DB |||||

QY 181 TGATATGAGCGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGCGGA 240
DB |||||

QY 241 ACATCATGTATGACGGGAAACACATGTACAGAAATGTCTTCCACCGGACGAGGCGTC 300
DB |||||

QY 241 ACATCATGTATGACGGGAAACACATGTACAGAAATGTCTTCCACCGGACGAGGCGTC 300
DB |||||

QY 301 CCGAGGACGAGGAGACATGAGGACGACCTTCAGACCATCTTTAGGCCCCCAAAATC 360
DB |||||

QY 301 CCGAGGACGAGGAGACATGAGGACGACCTTCAGACCATCTTTAGGCCCCCAAAATC 360
DB |||||

QY 361 TGAGGCTGTCTCACTCCAGCGCTCTGTGCAATATCAATATGAACCTTCCAAAGTGCCC 420
DB |||||

QY 361 TGAGGCTGTCTCACTCCAGCGCTCTGTGCAATATCAATATGAACCTTCCAAAGTGCCC 420
DB |||||

QY 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACAGACTTTG 480
DB |||||

QY 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACAGACTTTG 480
DB |||||

QY 481 TACCTTTCCCCCAACCGATCGCTTCAGCGCAAGGCCCTTCCCTCCCTCGCCCAATCA 540
DB |||||

QY 481 TACCTTTCCCCCAACCGATCGCTTCAGCGCAAGGCCCTTCTCCCTCCCTCGCCCAATCA 540
DB |||||

QY 541 GGCAGCTTTTCCCAACCGATCGCTTCAGCGCAAGGCCCTTCCCTCGCTTTTCCCTC 600
DB |||||

QY 541 GGCAGCTTTTCCCAACCGATCGCTTCAGCGCAAGGCCCTTCCCTCGCTTTTCCCTC 600
DB |||||

QY 601 CCATGCCACCAACCAATGCTTTGCTTAATAACCCCGAGTCCCTGGGCGACCTCTGGAC 660
DB |||||

DB 601 CCATGCCACCAACAATGCTTGTCTTAATAACCCCGAGTCCCTGGGGCACCTCTCTGGAC 660
QY |||||

DB 661 AAGGCACCTTTCCCTTCATGATGCCCTCCCTCCATGCTCATCCCGCGCCCTCCAG 720
QY |||||

DB 661 AAGGCACCTTTCCCTTCATGATGCCCTCCCTCCATGCTCATCCCGCGCCCTCCAG 720
QY |||||

DB 721 TCATGCCGAGCAGGTTAATTAATCAAGTACCTTCGGGCTTATTTCAACCAAACTTCCCAC 780
QY |||||

DB 721 TCATGCCGAGCAGGTTAATTAATCAAGTACCTTCGGGCTTATTTCTACCAAACTTCCCAC 780
QY |||||

DB 781 CTCCTCAGTTTTAATAGTTTCCAGAACCACTTAGTTCTTCTGCTGCTGCTTAATAACA 840
QY |||||

DB 781 CTCCTCAGTTTTAATAGTTTCCAGAACCACTTAGTTCTTCTGCTGCTGCTTAATAACA 840
QY |||||

DB 841 GCAGTAGTCTCTCAATTTACAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAAGTGA 900
QY |||||

DB 841 GCAGTAGTCTCTCAATTTACAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAAGTGA 900
QY |||||

DB 901 GAAGTCTCCAGAAAGGCTGAACAATATGATGACACAGGACCCAGACCAAGTCATG 960
QY |||||

DB 901 GAAGTCTCCAGAAAGGCTGAACAATATGATGACACAGGACCCAGACCAAGTCATG 960
QY |||||

DB 961 GCGAGGTGAGAGGATCGGTCCTGCTGATCGGCGGAGCGAGCCGCTCCGACAGA 1020
QY |||||

DB 961 GCGAGGTGAGAGGATCGGTCCTGCTGATCGGCGGAGCGAGCCGCTCCGACAGA 1020
QY |||||

DB 1021 GAAGCAAGACACGCGGTACAGATCTGATTATGACCGAGGGAGAACACCATCTCGGCACC 1080
QY |||||

DB 1021 GAAGCAAGACACGCGGTACAGATCTGATTATGACCGAGGGAGAACACCATCTCGGCACC 1080
QY |||||

DB 1081 GCAGCTTACGAAACGAGCAGAGACGAGAACCGGGAGACACAGGCATCGAACCAACCGAA 1140
QY |||||

DB 1081 GCAGCTTACGAAACGAGCAGAGACGAGAACCGGGAGACACAGGCATCGAACCAACCGAA 1140
QY |||||

DB 1141 GATCACCATCTCTGAAAGGCTCTTCAAAAAGAGTATGAAGATCTGAGAGGTTACG 1200
QY |||||

DB 1141 GATCACCATCTCTGAAAGGCTCTTCAAAAAGAGTATGAAGATCTGAGAGGTTACG 1200
QY |||||

DB 1201 GTTTATCGGTGTCTCTGAACCTGTGGATGCACACAGAAATTACCTGGGGAGATTATTA 1260
QY |||||

DB 1201 GTTTATCGGTGTCTCTGAACCTGTGGATGCACACAGAAATTACCTGGGGAGATTATTA 1260
QY |||||

DB 1261 AAAATAACAGATCTTTGGGCCCAACCTCTGGAGATTGTGAATCATCTCTCCCAAGTAGG 1320
QY |||||

DB 1261 AAAATAACAGATCTTTGGGCCCAACCTCTGGAGATTGTGAATCATCTCTCCCAAGTAGG 1320
QY |||||

DB 1321 AGAAGAGAGAGCTCTGTTGGGAGGAGAAAAGACCGTTGGAGTGACACCAAGTTCTG 1380
QY |||||

DB 1321 AGAAGAGAGAGCTCTGTTGGGAGGAGAAAAGACCGTTGGAGTGACACCAAGTTCTG 1380
QY |||||

DB 1381 GCAAGACAAAGAACATATACCTCAATCAAGGAAAAGAGCCCGAGGAGACCATGCTTGACA 1440
QY |||||

DB 1381 GCAAGACAAAGAACATATACCTCAATCAAGGAAAAGAGCCCGAGGAGACCATGCTTGACA 1440
QY |||||

DB 1441 AGAATGAGGAGGAGAAAGAAAGAACTTTTAAAGCTGTGTGGATTCCATGCATCTATTAG 1500
QY |||||

DB 1441 AGAATGAGGAGGAGAAAGAAAGAACTTTTAAAGCTGTGTGGATTCCATGCATCTATTAG 1500
QY |||||

DB 1501 AAAAATTAATCTTCCAGTGACCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA 1560
QY |||||

DB 1501 AAAAATTAATCTTCCAGTGACCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA 1560
QY |||||

DB 1561 GTAGGCTTCTGACTTATATGACAAATTTGAGAGGAGTTGGGAGCAGGCAAGAAAGG 1620
QY |||||

DB 1561 GTAGGCTTCTGACTTATATGACAAATTTGAGAGGAGTTGGGAGCAGGCAAGAAAGG 1620
QY |||||

DB 1621 CCAAGAGCTCTCGGCTCCGTTGGGAACTTCCAAAGACGAAGCTCGATGAAGAATTAGAGA 1680
QY |||||

DB 1621 CCAAGAGCTCTCGGCTCCGTTGGGAACTTCCAAAGACGAAGCTCGATGAAGAATTAGAGA 1680
QY |||||

DB 1681 GTTTCAGTGATCTCGAGTGTGAGTGTGATGAGGACGACCTGTTCTAGGAGCTCAGACT 1740
QY |||||

DB 1681 GTTTCAGTGATCTCGAGTGTGAGTGTGATGAGGACGACCTGTTCTAGGAGCTCAGACT 1740
QY |||||

QY	1741	CTGAACTTTTGGACGTTATTGAGAAATCAAACGCAAAAAGGCCACCTTGACCGACTTC	1800
DB	1741		
QY	1801	ATGATCAACTTTGGTCAACAGATCCAGGCGCAGATGAATGATGACCACTCTCGCAATGCA	1860
DB	1801		
QY	1861	ATGATGAACTTTGGTACACAGATCCAGGCGCAGATGAATGATGACCACTCTCGCAATGCA	1860
DB	1861		
QY	1861	GGCAAAAGCAAGACGACACAGGAAATAGGCACAGCAATTTATCTGTGAGAGAGGCCATCA	1920
DB	1861		
QY	1921	AGCCCTGTCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCAGAGTCT	1980
DB	1921	AGCCCTGTCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCAGAGTCT	1980
QY	1981	CCCCGCTACGAACTTTTAACTGACAGCGCAACTGTTATAGAAATACGATGATCAGAGT	2040
DB	1981	CCCCGCTACGAACTTTTAACTGACAGCGCAACTGTTATAGAAATACGATGATCAGAGT	2040
QY	2041	ATATCTTTGAAGGATTTTCTATGTTGCAATGCCCCCTGACCAATATTCACATGTGTA	2100
DB	2041	ATATCTTTGAAGGATTTTCTATGTTGCAATGCCCCCTGACCAATATTCACATGTGTA	2100
QY	2101	AAGTAATAGATTCAACATAGACTACAGATTCAATTTCAATGGAAGATGATGCCGGAGA	2160
DB	2101	AAGTAATAGATTCAACATAGACTACAGATTCAATTTCAATGGAAGATGATGCCGGAGA	2160
QY	2161	ATTTTGTGCGAAAGGCTTGAACCTTTTCACTGTTCTCTATTTCAGAGATATTTTGGAAAT	2220
DB	2161	ATTTTGTGCGAAAGGCTTGAACCTTTTCACTGTTCTCTATTTCAGAGATATTTTGGAAAT	2220
QY	2221	TATATGACTGGAATCTTAAAGGTCTTTGTTTGAAGACAGCCCTCCCTGCTGCCAAGAT	2280
DB	2221	TATATGACTGGAATCTTAAAGGTCTTTGTTTGAAGACAGCCCTCCCTGCTGCCAAGAT	2280
QY	2281	TTCAATTTGATGCAAGTTTGAAGATTTCTTCCAGATGAGAGAAAGGAAGTCTGTCCA	2340
DB	2281	TTCAATTTGATGCAAGTTTGAAGATTTCTTCCAGATGAGAGAAAGGAAGTCTGTCCA	2340
QY	2341	TGCACAGATTTCTCTGTACTTTTAAAGGTGAGCAAAAGCCCTGTGCTGAGGAGGAGA	2400
DB	2341	TGCACAGATTTCTCTGTACTTTTAAAGGTGAGCAAAAGCCCTGTGCTGAGGAGGAGA	2400
QY	2401	TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAG	2460
DB	2401	TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAG	2460
QY	2461	GCAATGATTTTACCAACCTCTGGGACGAAACCAAGCTCTGTCTCGTATCGATCAACTGGATC	2520
DB	2461	GCAATGATTTTACCAACCTCTGGGACGAAACCAAGCTCTGTCTCGTATCGATCAACTGGATC	2520
QY	2521	GTGAAACAGTTCAACCCCGATGTGATTTCTTTCCGATTTATCGTCCACTTTGGGATACGCC	2580
DB	2521	GTGAAACAGTTCAACCCCGATGTGATTTCTTTCCGATTTATCGTCCACTTTGGGATACGCC	2580
QY	2581	CTGCAAGTTGAGTTATGACGAGAGCCACAGTACCAAAAACTGTGGAAAGATTATGTA	2640
DB	2581	CTGCAAGTTGAGTTATGACGAGAGCCACAGTACCAAAAACTGTGGAAAGATTATGTA	2640
QY	2641	AACCTTGGCCACTCTCTAGCAAAATAGTCCCAAAAGTCAAACTGACAAAAGAGCTGG	2700
DB	2641	AACCTTGGCCACTCTCTAGCAAAATAGTCCCAAAAGTCAAACTGACAAAAGAGCTGG	2700
QY	2701	CACAGAGGAGGAGGAGCCCTTCAAAAATACGCGAGAGATACATGACAGAGATTA	2760
DB	2701	CACAGAGGAGGAGGAGCCCTTCAAAAATACGCGAGAGATACATGACAGAGATTA	2760
QY	2761	CGGTGAGCTAAGTAGCAAGATTTCTGAAACTGCGATCCGTTCTGATGCTGTGACG	2820
DB	2761	CGGTGAGCTAAGTAGCAAGATTTCTGAAACTGCGATCCGTTCTGATGCTGTGACG	2820

QY	2821	ATGCAATGATGCTACTGTTCTGACCCATCATATCGCTACCAACCAATGCTTAATGCAATT	2880
DB	2821	ATGCAATGATGCTACTGTTCTGACCCATCATATCGCTACCAACCAATGCTTAATGCAATT	2880
QY	2881	TGGACAAGTTGATAGGATATACCTTTCCAAAGATCGTTGTCTGTGACAGCTGGCCATGACTC	2940
DB	2881	TGGACAAGTTGATAGGATATACCTTTCCAAAGATCGTTGTCTGTGACAGCTGGCCATGACTC	2940
QY	2941	ATCCAAGTCACTATTTAAATTTTGGAAATGAAATCTCTGATCATGCGCAAGATTCATTAATCTA	3000
DB	2941	ATCCAAGTCACTATTTAAATTTTGGAAATGAAATCTCTGATCATGCGCAAGATTCATTAATCTA	3000
QY	3001	ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAGTTTCATCATGCAATCGGCA	3060
DB	3001	ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAGTTTCATCATGCAATCGGCA	3060
QY	3061	AGAAAAGGATTAACACCTTTGATAAATATCATGTACAGCCCTTGGCCAAAGATGACCAACTC	3120
DB	3061	AGAAAAGGATTAACACCTTTGATAAATATCATGTACAGCCCTTGGCCAAAGATGACCAACTC	3120
QY	3121	CCTCGAGGATTAACCAATGAAACGGTTGGAAATTTCTCTGGGTGATGCTGTTGTTGAATTC	3180
DB	3121	CCTCGAGGATTAACCAATGAAACGGTTGGAAATTTCTCTGGGTGATGCTGTTGTTGAATTC	3180
QY	3181	TGACCAAGCTCCATTTGTAATTTGTTTCTAGTCTGGAAGAGGAGATTTAGCAACT	3240
DB	3181	TGACCAAGCTCCATTTGTAATTTGTTTCTAGTCTGGAAGAGGAGATTTAGCAACT	3240
QY	3241	ATCGGACTGCCATTTGTTAGAAATCAGACCTTGGCCATGCTAGCAAAAGAACTTTGAACCTGG	3300
DB	3241	ATCGGACTGCCATTTGTTAGAAATCAGACCTTGGCCATGCTAGCAAAAGAACTTTGAACCTGG	3300
QY	3301	ATCCAATTAATGCTGATGCTCACGGCCCTGACCTTTGTGTAAGAAATCGGACCTTCGACATG	3360
DB	3301	ATCCAATTAATGCTGATGCTCACGGCCCTGACCTTTGTGTAAGAAATCGGACCTTCGACATG	3360
QY	3361	CAATCGCCAAATTTGTTGAAGGTTAATAGGAGCTGTTTACTTTGGAGGAAAGCTTGGAGG	3420
DB	3361	CAATCGCCAAATTTGTTGAAGGTTAATAGGAGCTGTTTACTTTGGAGGAAAGCTTGGAGG	3420
QY	3421	AAGCCAAGCAGTTATTTGGACCTTGTCTTTAAATGATCCGGACCTTGGCGAAGTCTGGC	3480
DB	3421	AAGCCAAGCAGTTATTTGGACCTTGTCTTTAAATGATCCGGACCTTGGCGAAGTCTGGC	3480
QY	3481	TCAATTAATCTCTCCACCACTCCAACTCAAGAGCCAAATATCTGATCGCAAACTTTATTTG	3540
DB	3481	TCAATTAATCTCTCCACCACTCCAACTCAAGAGCCAAATATCTGATCGCAAACTTTATTTG	3540
QY	3541	AAACTTCTCCAGTTCTCAAAAACCTTACTGAGTTTGAAGAAAGCAATTTGGAGTAAATTTTA	3600
DB	3541	AAACTTCTCCAGTTCTCAAAAACCTTACTGAGTTTGAAGAAAGCAATTTGGAGTAAATTTTA	3600
QY	3601	CTCATGTTTCGACTTCTGCGCAAGGCAATTCACATTCAGAACTCTGGGATTTAAACCACTCA	3660
DB	3601	CTCATGTTTCGACTTCTGCGCAAGGCAATTCACATTCAGAACTCTGGGATTTAAACCACTCA	3660
QY	3661	CCCTAGGCCACCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAGCCA	3720
DB	3661	CCCTAGGCCACCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAGCCA	3720
QY	3721	CAGAGTACTTATTTCAATTTCCAGATCATCATGAAAGGACACTTAACTTTGTTGCGAA	3780
DB	3721	CAGAGTACTTATTTCAATTTCCAGATCATCATGAAAGGACACTTAACTTTGTTGCGAA	3780
QY	3781	GCTCTTTGTTGTAATAATAGAACTCAGGCCAAGGTAGCGAGGAGCTGGGCATGCAAGGAT	3840
DB	3781	GCTCTTTGTTGTAATAATAGAACTCAGGCCAAGGTAGCGAGGAGCTGGGCATGCAAGGAT	3840
QY	3841	AGGCCAATAACCAAGCAAGACCAAGAGCCCTGTGGCGCTTCGCAACCAAGACTTTGGCGG	3900
DB	3841	AGGCCAATAACCAAGCAAGACCAAGAGCCCTGTGGCGCTTCGCAACCAAGACTTTGGCGG	3900
QY	3901	ACCTTTTGGAAATCATTTTATTTGACGCGCTGTACACTGATAAAGGATTTGGGAATATGTTTCA	3960

Db 3901 |||||ACCTTTGGAATCATTTATTGTCAGGCGCTGACACATGATGAGGATTTGGAAATATGTTTCATA 3960
Qy 3961 CTTTCATGAATGTCGCTCTCTTTCCACGATTAAGAAATTCATTTTGAATCAGGATTTGA 4020
Db 3961 CTTTCATGAATGTCGCTCTCTTTCCACGATTAAGAAATTCATTTTGAATCAGGATTTGA 4020
Qy 4021 ATGACCCCAAAATCCCAAGCTTCAGCAGTGTGCTTGACACCTTAGGACAGAAAGAGC 4080
Db 4021 ATGACCCCAAAATCCCAAGCTTCAGCAGTGTGCTTGACACCTTAGGACAGAAAGAGC 4080
Qy 4081 CAGACATTCCTCTGTACAAAGACTCTGACAGAGTGGGCCCATCCCATGCCGAACTTACA 4140
Db 4081 CAGACATTCCTCTGTACAAAGACTCTGACAGAGTGGGCCCATCCCATGCCGAACTTACA 4140
Qy 4141 CTGTGGCTGTTTATTTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGCAAGATATTTCAGC 4200
Db 4141 CTGTGGCTGTTTATTTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGCAAGATATTTCAGC 4200
Qy 4201 AAGCGAAATGGGAGCAGCAATGCGATGCGCTTGAAAAATATAATTTTCCCGAGATGGCC 4260
Db 4201 AAGCGAAATGGGAGCAGCAATGCGATGCGCTTGAAAAATATAATTTTCCCGAGATGGCC 4260
Qy 4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAAGAGTTAAAGAAATGAGTGGGAA 4320
Db 4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAAGAGTTAAAGAAATGAGTGGGAA 4320
Qy 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATTAAGAGGGCAT 4380
Db 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATTAAGAGGGCAT 4380
Qy 4381 GCAAGTGTGGAGTATTTACTTCTCAGTAACCTGACTGCTGCTATTTCAGACCTAGCCT 4440
Db 4381 GCAAGTGTGGAGTATTTACTTCTCAGTAACCTGACTGCTGCTATTTCAGACCTAGCCT 4440
Qy 4441 AGTTTTCTTCGACACAATGAACGAAGTGTGCTCAATTGAAATATAATACAGAGTCAAATCG 4500
Db 4441 AGTTTTCTTCGACACAATGAACGAAGTGTGCTCAATTGAAATATAATACAGAGTCAAATCG 4500
Qy 4501 CTATTGCTGTTTAAATGATCTGTTTTTAGCTGAGATGGTCTTTTATTAACAAGTATTAGATT 4560
Db 4501 CTATTGCTGTTTAAATGATCTGTTTTTAGCTGAGATGGTCTTTTATTAACAAGTATTAGATT 4560
Qy 4561 TTTCTTCTATTAAACGGAACCTTGACTTTGGTGAATGTGCATTACTCTCTTTTATTG 4620
Db 4561 TTTCTTCTATTAAACGGAACCTTGACTTTGGTGAATGTGCATTACTCTCTTTTATTG 4620
Qy 4621 CTCTTTAAATATAAAAATCAAGAAGCATATTCATGTGGAATAGATCCTGTTTTTCCAT 4680
Db 4621 CTCTTTAAATATAAAAATCAAGAAGCATATTCATGTGGAATAGATCCTGTTTTTCCAT 4680
Qy 4681 CTGTGTCAGATTTGTGACCTTAGACTTTCAATTTGACAAAGTAAATTTGACTTTACTAG 4740
Db 4681 CTGTGTCAGATTTGTGACCTTAGACTTTCAATTTGACAAAGTAAATTTGACTTTACTAG 4740
Qy 4741 TAAAAAATTT 4764
Db 4741 TAAAAAATTT 4764

RESULT 8
ADY17385
ID ADY17385 standard; DNA; 4764 BP.
XX
AC
ADY17385;
XX
DT 05-MAY-2005 (first entry)
DE
XX DNA encoding a PRO polypeptide, SEQ ID NO 3191.
XX Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antichyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;

KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
XX Antiallergic; ds; gene; diagnosis.
OS Homo sapiens.
XX WO2005016962-A2.
PN
XX 24-FEB-2005.
PD
XX 11-AUG-2004; 2004WO-US026249.
PF
XX 11-AUG-2003; 2003US-0493546P.
PR (GETH) GENENTECH INC.
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
PI WPI; 2005-182330/19.
DR
XX
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and treating an immune related disorder, e.g. systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
PS Claim 1; SEQ ID NO 3191; 158pp; English.
XX The invention relates to an isolated nucleic acid encoding a PRO polypeptide. The polypeptide, agonist or an antagonist, antibody, composition, and method are useful for diagnosing and treating an immune related disorder, e.g. systemic lupus erythematosus, rheumatoid arthritis. The present sequence represents a DNA encoding a PRO polypeptide.
XX
SQ Sequence 4764 BP; 1379 A; 1138 C; 1082 G; 1165 T; 0 U; 0 Other;

Query Match 99.9%; Score 4760.8; DB 14; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGCTCTGCTACTCGGTAGTAGCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG 60
Db 1 CTGCTCTGCTACTCGGTAGTAGCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG 60
Qy 61 AGCCTTTTATAGTTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
Db 61 AGCCTTTTATAGTTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
Qy 121 CCCATATCATCAAGAGCTGATAAATCTAGTGAAGAGTTAGACGTGTGCATCTTCACTA 180
Db 121 CCCATATCATCAAGAGCTGATAAATCTAGTGAAGAGTTAGACGTGTGCATCTTCACTA 180
Qy 181 TGATATGAGGCGAGTCTCTGAGCTTATTTCTCTGTGGAAGATGTGCATATCCAGGCGGA 240
Db 181 TGATATGAGGCGAGTCTCTGAGCTTATTTCTCTGTGGAAGATGTGCATATCCAGGCGGA 240
Qy 241 ACATCATGATGAGGGAACACATGTCACAGAAATGTCTTCCACCGGACGAGGCGTC 300
Db 241 ACATCATGATGAGGGAACACATGTCACAGAAATGTCTTCCACCGGACGAGGCGTC 300
Qy 301 CCCGAGACGAGGAGGACATGGAGCCAGACCTTCAGCACCATCTTTAGGCCCCCAATC 360
Db 301 CCCGAGACGAGGAGGACATGGAGCCAGACCTTCAGCACCATCTTTAGGCCCCCAATC 360
Qy 361 TGAGGCTGCTTCCACCTTCAGCAGCCTCTCTGTCATATCAATATGAACTTCCAAAGTGCCC 420
Db 361 TGAGGCTGCTTCCACCTTCAGCAGCCTCTCTGTCATATCAATATGAACTTCCAAAGTGCCC 420
Qy 421 CTTTCCACCACTTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACGAGCTTIG 480
Db 421 CTTTCCACCACTTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACGAGCTTIG 480
Qy 481 TACCTTCCCCCACCACCATGCTTCGTGACGAGGCGCTCTTCCCGCTGCCCCCAATCA 540
Db 481 TACCTTCCCCCACCACCATGCTTCGTGACGAGGCGCTCTTCCCGCTGCCCCCAATCA 540

QY	541	GGCGGCTTTTCCCAACCA	CAGATGAGGACCCCTTCC	CAGTTCTCTGTTTTCCTC	600
DB	541	GGCGGCTTTTCCCAACCA	CAGATGAGGACCCCTTCC	CAGTTCTCTGTTTTCCTC	600
QY	601	CCATGCCACCAACAAATG	CTTGTCTTAATAAACCCCA	GGTCCCTTGGGGCACCTCCT	660
DB	601	CCATGCCACCAACAAATG	CTTGTCTTAATAAACCCCA	GGTCCCTTGGGGCACCTCCT	660
QY	661	AAGGCACTTTCCCTTCAT	GATGAGCCCTCCCTCCAT	GTGCTCATGCCCGCCCTCCAG	720
DB	661	AAGGCACTTTCCCTTCAT	GATGAGCCCTCCCTCCAT	GTGCTCATGCCCGCCCTCCAG	720
QY	721	TCAATCCGACGAGGTAA	TATACAGTACCTCGGGCTA	TCTCACCACAATCTCCAC	780
DB	721	TCAATCCGACGAGGTAA	TATACAGTACCTCGGGCTA	TCTCACCACAATCTCCAC	780
QY	781	CTCCCACTTTTAATAGT	TTCCAGAAACCCCTAGT	TTCTTCTCGCCAGTGCTAA	840
DB	781	CTCCCACTTTTAATAGT	TTCCAGAAACCCCTAGT	TTCTTCTCGCCAGTGCTAA	840
QY	841	GCAGTAGTCTCTATTT	CAGACATCTCCCTCCATA	CCCACTCCCAAGGCTCCC	900
DB	841	GCAGTAGTCTCTATTT	CAGACATCTCCCTCCATA	CCCACTCCCAAGGCTCCC	900
QY	901	GAAGGTCCCAAGAAAG	CTGAAACATATGATGAC	CAACAGGCAACCGACCA	960
DB	901	GAAGGTCCCAAGAAAG	CTGAAACATATGATGAC	CAACAGGCAACCGACCA	960
QY	961	GGCGAGGTGAGGCA	TCGGTCCCTGGATCGG	CGGGAGCGAGGCCCGG	1020
DB	961	GGCGAGGTGAGGCA	TCGGTCCCTGGATCGG	CGGGAGCGAGGCCCGG	1020
QY	1021	GAAGACAAGACAGC	CGGTATCAGATCTGAT	TATGACCGAGGGAGAAC	1080
DB	1021	GAAGACAAGACAGC	CGGTATCAGATCTGAT	TATGACCGAGGGAGAAC	1080
QY	1081	GCAGCTACGAAACG	GAGCGAGAACCGGAGAG	ACACAGGCTCGAGACAC	1140
DB	1081	GCAGCTACGAAACG	GAGCGAGAACCGGAGAG	ACACAGGCTCGAGACAC	1140
QY	1141	GATCACCATCTCTG	GAAGGCTCTACAAAAG	AGAGTATAGAGATCTG	1200
DB	1141	GATCACCATCTCTG	GAAGGCTCTACAAAAG	AGAGTATAGAGATCTG	1200
QY	1201	GTTTATCGGTTGTT	CTCTGAACTCTGATG	CAACAGAAATTAACCTG	1260
DB	1201	GTTTATCGGTTGTT	CTCTGAACTCTGATG	CAACAGAAATTAACCTG	1260
QY	1261	AAATAACAGATTCT	TGGGCCCCACCCCTG	GAGATTGTGAATCATG	1320
DB	1261	AAATAACAGATTCT	TGGGCCCCACCCCTG	GAGATTGTGAATCATG	1320
QY	1321	AGAAAGAGAGCTCT	TGGGAGAGAAAGAC	CGTTGGATGACACCCAG	1380
DB	1321	AGAAAGAGAGCTCT	TGGGAGAGAAAGAC	CGTTGGATGACACCCAG	1380
QY	1381	GCAAGACAAGAACT	ATACCTCAATCAAGG	AAAAAGCCGAGGAGAC	1440
DB	1381	GCAAGACAAGAACT	ATACCTCAATCAAGG	AAAAAGCCGAGGAGAC	1440
QY	1441	AGAATCAGGAGGA	GAAGAACTTTTAAG	CCTGTGTGGAATTCG	1500
DB	1441	AGAATCAGGAGGA	GAAGAACTTTTAAG	CCTGTGTGGAATTCG	1500
QY	1501	AAACTACTACTCT	CAGTGACCCCATGAT	CAGGTGGAGATTCTA	1560
DB	1501	AAACTACTACTCT	CAGTGACCCCATGAT	CAGGTGGAGATTCTA	1560
QY	1561	GTAGGCTTCTGAT	TTATATGACAAATTT	TGAGGAGGTTGGG	1620
DB	1561	GTAGGCTTCTGAT	TTATATGACAAATTT	TGAGGAGGTTGGG	1620

QY	1621	CCAAAGCTGCTCG	CCCTCGTGGAAAC	CTCCAAAGACGAGCT	CGATGAGATTTAGAG	1680
DB	1621	CCAAAGCTGCTCG	CCCTCGTGGAAAC	CTCCAAAGACGAGCT	CGATGAGATTTAGAG	1680
QY	1681	GTTTCCAGTGAAT	CCGAGTGTGAGT	CTGATGAGACACG	CACTCTTCTAGCAGT	1740
DB	1681	GTTTCCAGTGAAT	CCGAGTGTGAGT	CTGATGAGACACG	CACTCTTCTAGCAGT	1740
QY	1741	CTGAAGTTTGTG	AGCTTATTTGAGAA	ATCAAACGCAAAA	AGGCCCACTTGAC	1800
DB	1741	CTGAAGTTTGTG	AGCTTATTTGAGAA	ATCAAACGCAAAA	AGGCCCACTTGAC	1800
QY	1801	ATGATCAACTTT	TGGTACACGAT	TCCAGGCCAGAT	GATGATGACCACT	1860
DB	1801	ATGATCAACTTT	TGGTACACGAT	TCCAGGCCAGAT	GATGATGACCACT	1860
QY	1861	GGCAAAAGCACA	GACGACAGGAAT	TAGGCACAGCAT	TTTATCTGGAGAGAG	1920
DB	1861	GGCAAAAGCACA	GACGACAGGAAT	TAGGCACAGCAT	TTTATCTGGAGAGAG	1920
QY	1921	AGCCCTGTCTG	TCCTATGACAA	CAATGCTGGCAG	ACTTTTCCACTAC	1980
DB	1921	AGCCCTGTCTG	TCCTATGACAA	CAATGCTGGCAG	ACTTTTCCACTAC	1980
QY	1981	CCCCGCTACG	AACTTTTAACTG	ACAGGCCAACT	GTGTATAGAAAT	2040
DB	1981	CCCCGCTACG	AACTTTTAACTG	ACAGGCCAACT	GTGTATAGAAAT	2040
QY	2041	ATATCTTTCA	AGGATTTCTATG	TTGTCATGCCCC	CTGACCAATAT	2100
DB	2041	ATATCTTTCA	AGGATTTCTATG	TTGTCATGCCCC	CTGACCAATAT	2100
QY	2101	AAGTAAATAG	ATTCAAATAG	ACTACACGAT	TTCATTTGAAGAG	2160
DB	2101	AAGTAAATAG	ATTCAAATAG	ACTACACGAT	TTCATTTGAAGAG	2160
QY	2161	ATTTTGTCTG	AAAGGCTTGA	ACTTTTCACTG	TTCTTATTCAGAG	2220
DB	2161	ATTTTGTCTG	AAAGGCTTGA	ACTTTTCACTG	TTCTTATTCAGAG	2220
QY	2221	TATATGACTG	GAATCTTAAAG	GTCTTTTGAAG	ACAGCCCTCCCT	2280
DB	2221	TATATGACTG	GAATCTTAAAG	GTCTTTTGAAG	ACAGCCCTCCCT	2280
QY	2281	TTCAATTCAT	CGCCAGTTTGT	TAAGATTTCTT	CCAGATCGAGAA	2340
DB	2281	TTCAATTCAT	CGCCAGTTTGT	TAAGATTTCTT	CCAGATCGAGAA	2340
QY	2341	TGCAACAGAT	TTCTCTGTACT	TTAAGGTGACG	AAAGCCCTGGT	2400
DB	2341	TGCAACAGAT	TTCTCTGTACT	TTAAGGTGACG	AAAGCCCTGGT	2400
QY	2401	TTGGCAATAT	CTTCAAGTGG	AGAGCTGAGT	GCAGAAATATG	2460
DB	2401	TTGGCAATAT	CTTCAAGTGG	AGAGCTGAGT	GCAGAAATATG	2460
QY	2461	GCATGATTCT	TACCAACCCCT	GGGACGAAAC	CAAGCTCTG	2520
DB	2461	GCATGATTCT	TACCAACCCCT	GGGACGAAAC	CAAGCTCTG	2520
QY	2521	GTGAAACAG	TTCAACCCCG	ATGTACTTTT	CCGATATCGT	2580
DB	2521	GTGAAACAG	TTCAACCCCG	ATGTACTTTT	CCGATATCGT	2580
QY	2581	CTGCAAGTTG	ATGATGAGG	AGAGCCACAG	TATACCAAACT	2640
DB	2581	CTGCAAGTTG	ATGATGAGG	AGAGCCACAG	TATACCAAACT	2640
QY	2641	AACTTCGCC	CACTCTCTAG	CAAAATAGT	TCCCAAACTG	2700
DB	2641	AACTTCGCC	CACTCTCTAG	CAAAATAGT	TCCCAAACTG	2700
QY	2701	CAAGAGGG	AGGAAAGCCCT	CTCCAAAAAT	ATACGGCAGAA	2760

Db 2701 |||||CACAGAGGAGGAGCCCTCCAAAAATACGGCAGAGAATACAAATGAGACGAGAGTAA 2760
Qy 2761 CGGTGAGCTAGTAGCAAGGATCTTGAAAACTGGCATCCGTTCTGTGATCTGTGCAGC 2820
Db 2761 CGGTGAGCTAGTAGCAAGGATCTTGAAAACTGGCATCCGTTCTGTGATCTGTGCAGC 2820
Qy 2821 ATGCATGATGCTACCTGTTCTGACCCATCATATCCGCTACCACTAAATGCCATTCAT 2880
Db 2821 ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCACTAAATGCCATTCAT 2880
Qy 2881 TGGACAAGTTGATGATGATATACCTTCCAAAGATCGTTGTCTGTGACAGCTGGCCATGACTC 2940
Db 2881 TGGACAAGTTGATGATGATATACCTTCCAAAGATCGTTGTCTGTGACAGCTGGCCATGACTC 2940
Qy 2941 ATCCAAAGTCAATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGACGAAATTCATATCTA 3000
Db 2941 ATCCAAAGTCAATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGACGAAATTCATATCTA 3000
Qy 3001 ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAGAAATTCATCATGACATGCATCGCGA 3060
Db 3001 ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAGAAATTCATCATGACATGCATCGCGA 3060
Qy 3061 AGAAAGGATTAACACTTGTGATAAATATCATGTACGCTTGGCCAAAGATGACCCAACTC 3120
Db 3061 AGAAAGGATTAACACTTGTGATAAATATCATGTACGCTTGGCCAAAGATGACCCAACTC 3120
Qy 3121 CCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTGTTGAATTC 3180
Db 3121 CCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTGTTGAATTC 3180
Qy 3181 TGAACGAGCTCCATTTGTACTATTTGTTCTTCTAGTCTGGAAGAGAGGATTAGCAACT 3240
Db 3181 TGAACGAGCTCCATTTGTACTATTTGTTCTTCTAGTCTGGAAGAGAGGATTAGCAACT 3240
Qy 3241 ATCGGACTGCATTTGTTCAAGATCAGACCTTGCATGCTAGCAAGAACTTGAACCTG 3300
Db 3241 ATCGGACTGCATTTGTTCAAGATCAGACCTTGCATGCTAGCAAGAACTTGAACCTG 3300
Qy 3301 ATCCATTTATGCTGATGCTACAGGCGCTGACCTTTGTAGAGAACTCGGACCTTCGACATG 3360
Db 3301 ATCCATTTATGCTGATGCTACAGGCGCTGACCTTTGTAGAGAACTCGGACCTTCGACATG 3360
Qy 3361 CAATGCGCAATTTGTTGAAGCTTAATAGGAGCTGTTTACTTTGAGGAGGAGGCTCGAGG 3420
Db 3361 CAATGCGCAATTTGTTGAAGCTTAATAGGAGCTGTTTACTTTGAGGAGGAGGCTCGAGG 3420
Qy 3421 AAGCCAAAGCAGTTATTTGGACGCTTGTCTTTTAATGATCCGACCTCGCGAAGTCTGGC 3480
Db 3421 AAGCCAAAGCAGTTATTTGGACGCTTGTCTTTTAATGATCCGACCTCGCGAAGTCTGGC 3480
Qy 3481 TCAATTTATCCTCTCCACCACTCCAACTACAAGAGCCAAATCTGATGACAACTTATTTG 3540
Db 3481 TCAATTTATCCTCTCCACCACTCCAACTACAAGAGCCAAATCTGATGACAACTTATTTG 3540
Qy 3541 AAACCTTCTCCAGTTCTACAAAATTTACTGAGTTTGAAGCAATTTGGAGTAAATTTTA 3600
Db 3541 AAACCTTCTCCAGTTCTACAAAATTTACTGAGTTTGAAGCAATTTGGAGTAAATTTTA 3600
Qy 3601 CTCATGTTGCACTTCTGGCAAGGCAATTCATTTGAGAACTGTTGGGATTTAAACCATCTGA 3660
Db 3601 CTCATGTTGCACTTCTGGCAAGGCAATTCATTTGAGAACTGTTGGGATTTAAACCATCTGA 3660
Qy 3661 CCTAGGCCCAATACAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAGCCA 3720
Db 3661 CCTAGGCCCAATACAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAGCCA 3720
Qy 3721 CAGAGTACTTATTCATTTCCAGATCATCATGAGGACACTTAACTTTGTTGCGAA 3780
Db 3721 CAGAGTACTTATTCATTTCCAGATCATCATGAGGACACTTAACTTTGTTGCGAA 3780
Qy 3781 GCTCTTTTGGTGAATATAGAACTCAGGCCAAGTACGCGAGGAGCTGGGCATGACGAGT 3840

Db 3781 GCTCTTTTGGTGAATATAGAACTCAGGCCAAGGTAGCGAGGAGCTGGGCATGCAGAGT 3840
Qy 3841 AGCCCATTAACCAACGACAGACCAAGAGCCCTGTGCGCTTCGCAACCAAGACCTTGGCGG 3900
Db 3841 AGCCCATTAACCAACGACAGACCAAGAGCCCTGTGCGCTTCGCAACCAAGACCTTGGCGG 3900
Qy 3901 ACCTTTTGGAAATCATTTATTTGACGCGCTGTACACTGATAAGGATTTGGAAATATGTTTATA 3960
Db 3901 ACCTTTTGGAAATCATTTATTTGACGCGCTGTACACTGATAAGGATTTGGAAATATGTTTATA 3960
Qy 3961 CTTTCATGAATGCTGCTCTTTTCCAGATTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
Db 3961 CTTTCATGAATGCTGCTCTTTTCCACGATTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
Qy 4021 ATGACCCCAAAATCCAGCTTTCAGAGCTGCTGACACTTATAGGACAGAAAGAGAGC 4080
Db 4021 ATGACCCCAAAATCCAGCTTTCAGAGCTGCTGACACTTATAGGACAGAAAGAGAGC 4080
Qy 4081 CAGACATTCCTCTGTACAGACTCTGCGACAGCTGCGGCCCATCCCATGCCCCGAACCTACA 4140
Db 4081 CAGACATTCCTCTGTACAGACTCTGCGACAGCTGCGGCCCATCCCATGCCCCGAACCTACA 4140
Qy 4141 CTGTGGCTGTTTATTTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGACCAAGTATTCAGC 4200
Db 4141 CTGTGGCTGTTTATTTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGACCAAGTATTCAGC 4200
Qy 4201 AAGCGGAAATGGGACGCAATGGATGCGCTTGGAAAAATATATTTTCCCCAGATGGCCC 4260
Db 4201 AAGCGGAAATGGGACGCAATGGATGCGCTTGGAAAAATATATTTTCCCCAGATGGCCC 4260
Qy 4261 ATCAGAAGCGGTTTCATCGAAACGAAAGTACAGACAGAGTTTAAAAAGAAATGAGTGGGAAA 4320
Db 4261 ATCAGAAGCGGTTTCATCGAAACGAAAGTACAGACAGAGTTTAAAAAGAAATGAGTGGGAAA 4320
Qy 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGGAAGACATCAAGAAATAAAGAGGAGGCAT 4380
Db 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGGAAGACATCAAGAAATAAAGAGGAGGCAT 4380
Qy 4381 GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACTGCTCTATTCAGACCTTAGCCT 4440
Db 4381 GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACTGCTCTATTCAGACCTTAGCCT 4440
Qy 4441 AGTTTTTCTCGACAGACAAATGAAGAGAGTGTGCTCATTTGAAATAAAAATACAGAGTCAAAATCG 4500
Db 4441 AGTTTTTCTCGACAGACAAATGAAGAGAGTGTGCTCATTTGAAATAAAAATACAGAGTCAAAATCG 4500
Qy 4501 CTATTTGTTGTTTAAATGATCTGTTTATAGCTGGAATGCTCTTTATTAACAAGTATTAGATT 4560
Db 4501 CTATTTGTTGTTTAAATGATCTGTTTATAGCTGGAATGCTCTTTATTAACAAGTATTAGATT 4560
Qy 4561 TTTCTCTATTTAAACGGAACCTGACTTTTGGTGAATGTGCAATCTCTCTTTTATTTTG 4620
Db 4561 TTTCTCTATTTAAACGGAACCTGACTTTTGGTGAATGTGCAATCTCTCTTTTATTTTG 4620
Qy 4621 CTCCTTTAAATAATAAAATCAAGAGACATATTTCTATGTGGAATAGATCTCTGTTTTCAT 4680
Db 4621 CTCCTTTAAATAATAAAATCAAGAGACATATTTCTATGTGGAATAGATCTCTGTTTTCAT 4680
Qy 4681 CTGTGTCAGATTTGACCTTAGACTTTCAATTTGACAAAGTAAAAAATTTGACTTTTACTAG 4740
Db 4681 CTGTGTCAGATTTGACCTTAGACTTTCAATTTGACAAAGTAAAAAATTTGACTTTTACTAG 4740
Qy 4741 TAAAAAATAAAAAATAAAAAATAAAAA 4764
Db 4741 TAAAAAATAAAAAATAAAAAATAAAAA 4764

RESULT 9
ADQ22183
ID ADQ22183 standard; DNA; 5425 BP.
XX
XX ADQ22183;


```
DT 26-AUG-2004 (first entry)
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 5003.
DE soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW db.
XX Homo sapiens.
OS
XX WO2004048938-A2.
PN
XX 10-JUN-2004.
PD
XX
XX 26-NOV-2003; 2003WO-US038193.
PF
XX 26-NOV-2002; 2002US-0429739P.
PR
XX (PROT-) PROTEIN DESIGN LABS INC.
PA
XX Aziz N, Ginsburg WM, Zlotnik A;
PI
XX WPI; 2004-441208/41.
DR
XX Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX
XX Example 2; SEQ ID NO 5003; 210bp; English.
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cyostatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 5425 BP; 1583 A; 1257 C; 1234 G; 1349 T; 0 U; 2 Other;
SQ
Query Match 99.3%; Score 4732.4; DB 12; Length 5425;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 4751; Conservative 0; Mismatches 11; Indels 2; Gaps 1;
QY 1 CTGTCTTGGTACCTGCGGTAGTAGCCTGGCTTTGCTCTGACGCGCATCTCGCGGCCGAG 60
DB 101 CTGTCTTGGTACCTGCGGTAGTAGCCTGGCTTTGCTCTGACGCGCATCTCGCGGCCGAG 160
QY 61 AGCCTTTTATAGTGTCTTCCCGGGGATGGAAGNTACAGAAATGACTGTGAATCAA 120
DB 161 AGCCTTTTATAGTGTCTTCCCGGGGATGGAAGNTACAGAAATGACTGTGAATCAA 220
QY 121 CCATATATCAAGGAGCTGATAATCTAGTGAAGAGTTAGACGTGTGCATCTTCACTA 180
DB 221 CCATATATCAAGGAGCTGATAATCTAGTGAAGAGTTAGACGTGTGCATCTTCACTA 280
QY 181 TGATATGAGGCGAGTCTCTGAGCTTATATTTCTGTGGAAGATGTGACATATCCAGCGGA 240
DB 281 TGATATGAGGCGAGTCTCTGAGCTTATATTTCTGTGGAAGATGTGACATATCCAGCGGA 340
QY 241 ACATCATATGAGGGAACACATGTCAAGAAATGTGTTCCACCGGACGAGGCGTC 300
DB 341 ACATCATATGAGGGAACACATGTCAAGAAATGTGTTCCACCGGACGAGGCGTC 400
QY 301 CCGGAGGAGGAGGACATGAGCCAGACCTTCAGCACCATCTTTAGGCCCCCAATC 360
DB 401 CCGGAGGAGGAGGACATGAGCCAGACCTTCAGCACCATCTTTAGGCCCCCAATC 460
```

1441 AGAATGAGGAGGAAGAAGAACTTTCTTAAGCCTGTGTGGATTCGATGCACCTCATTTCA 1500
1541 AGAATGAGGAGGAAGAAGAACTTTCTTAAGCCTGTGTGGATTCGATGCACCTCATTTCA 1600
1501 AAAAATACTATCTCCAGTGAACCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAA 1560
1601 AAAAATACTATCTCCAGTGAACCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAA 1660
1561 GTAGGCTTGTGATCTATATGACAAATTTGAGGAGGAGTTGGGAGCAGCGCAAGAAAGG 1620
1661 GTAGGCTTGTGATCTATATGACAAATTTGAGGAGGAGTTGGGAGCAGCGCAAGAAAGG 1720
1621 CCAAGCTGCTCGGCCTCCGTGGGAACCTCCAAAGACGAAGCTCGATGAAGATTAGAGA 1680
1721 CCAAGCTGCTCGGCCTCCGTGGGAACCTCCAAAGACGAAGCTCGATGAAGATTAGAGA 1780
1681 GTTCCAGTGAATCCGAGTGTGATGTGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1740
1781 GTTCCAGTGAATCCGAGTGTGATGTGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1840
1741 CTGAGTGTGACGTTATTCAGAAATCAAACGGCAAAAAGGCCACCTGACCCGACTTC 1800
1841 CTGAGTGTGACGTTATTCAGAAATCAAACGGCAAAAAGGCCACCTGACCCGACTTC 1900
1801 ATGATGAATTTGGTACAAAGTCCAGGCGCAGATGAATGATGACCACTCTGCAAAATGCA 1860
1901 A--AGGAATTTGGTACAAAGTCCAGGCGCAGATGAATGATGACCACTCTGCAAAATGCA 1958
1861 GCGCAAGGCAAGACGCAAGGAAATTAGGCAAGCAATTTATCTGGAGGAAGGCCATCA 1920
1959 GCGCAAGGCAAGACGCAAGGAAATTAGGCAAGCAATTTATCTGGAGGAAGGCCATCA 2018
1921 AGCCCTGTGCTCTATGACCAATGCTGGGAGACTTTTCCACTACCCGATCAGACTCT 1980
2019 AGCCCTGTGCTCTATGACCAATGCTGGGAGACTTTTCCACTACCCGATCAGACTCT 2078
1981 CCCGCTTACGAACTTTTAACTGACAGGCAACTGTTATAGAAATACGATGATCAGAGT 2040
2079 CCCGCTTACGAACTTTTAACTGACAGGCAACTGTTATAGAAATACGATGATCAGAGT 2138
2041 ATATCTTTGAAGGATTTCTATGTTTGCATATGCCCCCTGACCAATATTCACATGTGTA 2100
2139 ATATCTTTGAAGGATTTCTATGTTTGCATATGCCCCCTGACCAATATTCACATGTGTA 2198
2101 AGATAATAGATTTCAATAGACTACACGATTCATTTGATGAGAGATGATGCGGAGA 2160
2199 AGATAATAGATTTCAATAGACTACACGATTCATTTGATGAGAGATGATGCGGAGA 2258
2161 ATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTCTTATTCAGAGATATTTTGGAA 2220
2259 ATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTCTTATTCAGAGATATTTTGGAA 2318
2221 TATATGACTGGAACTTTAAAGGTCCTTGTGTTGAAGACAGCCCTCCCTGCCCCAAGAT 2280
2319 TATATGACTGGAACTTTAAAGGTCCTTGTGTTGAAGACAGCCCTCCCTGCCCCAAGAT 2378
2281 TTAATTTCAATGCAAGCTTTTGAAGATTTCTTCAGATGAGGAAGAGAGTCTGTCCA 2340
2379 TTAATTTCAATGCAAGCTTTTGAAGATTTCTTCAGATGAGGAAGAGAGTCTGTCCA 2438
2341 TGCAACAGATTTCTCTGTACTTGTAAAGTGTGAGCAAGCCCTGTGTGCTGAGGAGAGA 2400
2439 TGCAACAGATTTCTCTGTACTTGTAAAGTGTGAGCAAGCCCTGTGTGCTGAGGAGAGA 2498
2401 TTGCCAATATGCTTCAGTGGGAGAGCTGGAGTGGCAGAAATATGAGGAAGAAATGCAAG 2460
2499 TTGCCAATATGCTTCAGTGGGAGAGCTGGAGTGGCAGAAATATGAGGAAGAAATGCAAG 2558
2461 GCATGATTTGTATCAACCCCTGGGACGAACCAAGCTCTGTGCTGATTCGATCAACTGATC 2520
2559 GCATGATTTGTATCAACCCCTGGGACGAACCAAGCTCTGTGCTGATTCGATCAACTGATC 2618
2521 GTGAACAGTTCAACCCCGATGTGATTTACTTTTCCGATTTATCGTCCACTTTTGGGATACGCG 2580

2619 GTGAACAGTTCAACCCCGATGTGATTTACTTTTCCGATTTATCGTCCACTTTGGGATACGCG 2678
2581 CTGCACAGTTGAGTTATTCAGGAGACCCACAGTACCAAAAACCTGTGGAGAGTTATGTGA 2640
2679 CTGCACAGTTGAGTTATTCAGGAGACCCACAGTACCAAAAACCTGTGGAGAGTTATGTGA 2738
2641 AACTTTCGCCACCTCTAGCAATAGTCCCAAAGTCAAACAACCTGACAAAGAGAGCTGG 2700
2739 AACTTTCGCCACCTCTAGCAATAGTCCCAAAGTCAAACAACCTGACAAAGAGAGCTGG 2798
2701 CACAGAGGAGGAAGCCCTTCAAAAATAACGCGAGAGAATAAATCAATGAGAGAGAGTTAA 2760
2799 CACAGAGGAGGAAGCCCTTCAAAAATAACGCGAGAGAATAAATCAATGAGAGAGAGTTAA 2858
2761 CGGTGGAGTANGTAGCCCAAGGATTTCTGGAAACCTGGCATCCGTTCTGATGCTGTGCAG 2820
2859 CGGTGGAGTANGTAGCCCAAGGATTTCTGGAAACCTGGCATCCGTTCTGATGCTGTGCAG 2918
2821 ATGCAATGATGCTACTGTTCTGACCCCATCATATCCGCTTACCACCAATGCTTAATGCAAT 2880
2919 ATGCAATGATGCTACTGTTCTGACCCCATCATATCCGCTTACCACCAATGCTTAATGCAAT 2978
2881 TGGACAAGTTGATAGGATATATCTTTCCAAGATCGTTGCTGTGTGAGCTGGCCATGACTC 2940
2979 TGGACAAGTTGATAGGATATATCTTTCCAAGATCGTTGCTGTGTGAGCTGGCCATGACTC 3038
2941 ATCCAAGTCATCATTTAAATTTTGAANTGAATCCCTGATCATGCCAGGATTCATTTATCTA 3000
3039 ATCCAAGTCATCATTTAAATTTTGAANTGAATCCCTGATCATGCCAGGATTCATTTATCTA 3098
3001 ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAAGTTTCATCACATGACATGCGGA 3060
3099 ACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAAGTTTCATCACATGACATGCGGA 3158
3061 AGAAAGGATTAACACCTTGTGATAAATATCATGTCAAGCTTGGCCCAAGATGACCCAACTC 3120
3159 AGAAAGGATTAACACCTTGTGATAAATATCATGTCAAGCTTGGCCCAAGATGACCCAACTC 3218
3121 CCTCGAGGATTAACACATGAAACGTTGGAAATTCCTGGTGTATGCTGTTGTGAATTTTC 3180
3219 CCTCGAGGATTAACACATGAAACGTTGGAAATTCCTGGTGTATGCTGTTGTGAATTTTC 3278
3181 TGACACAGCTCATTTGTACTATTTGTTTCTAGTCTGGAAGAAGAGGATTTAGCAACTC 3240
3279 TGACACAGCTCATTTGTACTATTTGTTTCTAGTCTGGAAGAAGAGGATTTAGCAACTC 3338
3241 ATCGGACTGCCATTTGTCAGAAATCAGACCTTTGCCATGCTAGCAAAAGAACTTTGAAC 3300
3339 ATCGGACTGCCATTTGTCAGAAATCAGACCTTTGCCATGCTAGCAAAAGAACTTTGAAC 3398
3301 ATCGAATTTATGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
3399 ATCGAATTTATGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3458
3361 CAATGCGCAATTTGTTTGAAGCGTTAAATAGGAGCTGTTTACTTTGAGGGAAGCTTGGAGG 3420
3459 CAATGCGCAATTTGTTTGAAGCGTTAAATAGGAGCTGTTTACTTTGAGGGAAGCTTGGAGG 3518
3421 AAGCCAAAGCAGTTATTTGGACGCTTGTCTCTTTAAATGATCCGACCTTCGCGAAGTCTGCG 3480
3519 AAGCCAAAGCAGTTATTTGGACGCTTGTCTCTTTAAATGATCCGACCTTCGCGAAGTCTGCG 3578
3481 TCAATTTATCTCTCCACCCACTCCAACTACAGAGCCAAATATCTGATCCCAACTTTATTG 3540
3579 TCAATTTATCTCTCCACCCACTCCAACTACAGAGCCAAATATCTGATCCCAACTTTATTG 3638
3541 AAACCTTCTCCAGTCTTACAAAAAATTTACTGAGTTTGAAGAAGCAATTTGGAGTAATTTTA 3600
3639 AAACCTTCTCCAGTCTTACAAAAAATTTACTGAGTTTGAAGAAGCAATTTGGAGTAATTTTA 3698
3601 CTCATGTTTCGACTTCTGGCAAGGCAATTCACATTTGAGAACTGTGGGATTTAACCATCTGA 3660

3699	CTCATGTTGCACTTCTGTGCAAGGGCAATTCAATTGAGAACTGTGGGAAATTAAACAATCTGA	3755
Qy	CCCTAGGCCACAAATCAGAGAAATGGAAATTCCTAGGTGACTCCATAATGCAAACTGGTAGCCA	3720
	CCCTAGGCCACAAATCAGAGAAATGGAAATTCCTAGGTGACTCCATAATGCAAACTGGTAGCCA	3818
	CCCTAGGCCACAAATCAGAGAAATGGAAATTCCTAGGTGACTCCATAATGCAAACTGGTAGCCA	3818
3721	CAGAGTACTTATATTCATTTCAATTTCCAGATCATATGAAAGGACACTTAATCTTTGTTGGAA	3780
3819	CAGAGTACTTATATTCATTTCAATTTCCAGATCATATGAAAGGACACTTAATCTTTGTTGGAA	3878
3781	GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATGCAAGAGT	3840
3879	GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATGCAAGAGT	3938
3841	ACGCCATAACCAACGACAGAACCAAGAGGCCCTGTGGCGCTTCGCACCAAGACCTTGGCGG	3900
3939	ACGCCATAACCAACGACAGAACCAAGAGGCCCTGTGGCGCTTCGCACCAAGACCTTGGCGG	3998
3901	ACCTTTTGGAAATCAATTTATTGACAGCGCTGTACATCTGATTAAGGATTTGGAAATATGTTTCA	3960
3999	ACCTTTTGGAAATCAATTTATTGACAGCGCTGTACATCTGATTAAGGATTTGGAAATATGTTTCA	4058
3961	CTTTTCATGAATGCTGCTCTCTTTCCACGATTTGAAGAAATTCATTTTCAATCAGATTTGGA	4020
4059	CTTTTCATGAATGCTGCTCTCTTTCCACGATTTGAAGAAATTCATTTTCAATCAGATTTGGA	4118
4021	ATGACCCCAATCCACGCTTCAGCAGATGTTGCTTGACACTTAGGACAGAAAGGAAAAGAGC	4080
4119	ATGACCCCAATCCACGCTTCAGCAGATGTTGCTTGACACTTAGGACAGAAAGGAAAAGAGC	4178
4081	CAGACATTCCTCTGTGTAACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTACA	4140
4179	CAGACATTCCTCTGTGTAACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTACA	4238
4141	CTGTGGCTGTTATTTTCAAGGGGAAGAAGATAGCTGTGGGAAAGGACCAAGATTTTCAGC	4200
4239	CTGTGGCTGTTATTTTCAAGGGGAAGAAGATAGCTGTGGGAAAGGACCAAGATTTTCAGC	4298
4201	AAGCGGAAATGGGACAGCAATGATGCTGTGAAAAATATAATTTTCCCCAGATGGCCC	4260
4299	AAGCGGAAATGGGACAGCAATGATGCTGTGAAAAATATAATTTTCCCCAGATGGCCC	4358
4261	ATCAGAAGCGGTTTCAATCGAATCGGAAGTACAGACAAGAGTTTAAAGAAATGAGTGGGAAA	4320
4359	ATCAGAAGCGGTTTCAATCGAATCGGAAGTACAGACAAGAGTTTAAAGAAATGAGTGGGAAA	4418
4321	GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGAGAGGCGAT	4380
4419	GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGAGAGGCGAT	4478
4381	GCAAGTGTGGAGTATTTACTTTGCTCAGTAACTGTGACTGTGTGCTATTGAGACCTAGCCT	4440
4479	GCAAGTGTGGAGTATTTACTTTGCTCAGTAACTGTGACTGTGTGCTATTGAGACCTAGCCT	4538
4441	AGTTTTCCTCGACAAATGAACGAAGTGTGCTCATTTGAAATAAAATACAGAGTCAAAATCG	4500
4539	AGTTTTCCTCGACAAATGAACGAAGTGTGCTCATTTGAAATAAAATACAGAGTCAAAATCG	4598
4501	CTATTGTTGTTTAAATGATCTGTTTTTTAGCTGGATGCTTTTATTATCAAAAGTATTAGATT	4560
4599	CTATTGTTGTTTAAATGATCTGTTTTTTAGCTGGATGCTTTTATTATCAAAAGTATTAGATT	4658
4561	TTTTCTCTATTTTAAACGGAAACTTGACTTTGGTGAATGTCATTTACTTCCTTTTATTTTTG	4620
4659	TTTTCTCTATTTTAAACGGAAACTTGACTTTGGTGAATGTCATTTACTTCCTTTTATTTTTG	4718
4621	CTCTTTTAAATAAATAATTCAGAAAGCATATTTCTATGTGGAAATAGATCTCTGTTTTTCCAT	4680
4719	CTCTTTTAAATAAATAATTCAGAAAGCATATTTCTATGTGGAAATAGATCTCTGTTTTTCCAT	4778
4681	CTGTGTCCTCAGATTGTGACCTAGACTTTTCAATTTGACAAGTAAATAATTTGACTTTACTAG	4740
4779	CTGTGTCCTCAGATTGTGACCTAGACTTTTCAATTTGACAAGTAAATAATTTGACTTTACTAG	4838


```
QY 2484 ACCGAACCAAGCTCTGTCCTGATCGATCAACTGGATCGTGAAACAGTTTCAACCCCGATGTG 2543
Db 1717 ACGAACCAAGCTCTGTCCTGATCGATCAACTGGATCGTGAAACAGTTTCAACCCCGATGTG 1776
QY 2544 ATTACTTTTCCGATTATCGTCCACTTTGGGATACGCCCTGCACAGTTGAGTTATGCAGGA 2603
Db 1777 ATTACTTTTCCGATTATCGTCCACTTTGGGATACGCCCTGCACAGTTGAGTTATGCAGGA 1836
QY 2604 GACCCACAGTACCAGAAACCTGTGGAGAGAGTTATGTGAAACTTTCGCCACCTCTCTAGCAAT 2663
Db 1837 GACCCACAGTACCAGAAACCTGTGGAGAGAGTTATGTGAAACTTTCGCCACCTCTCTAGCAAT 1896
QY 2664 AGTCCCAAGTCAAAACCTGTGCAACAGAGCTGCAAGAGGGGAGGAGCCCTCCAA 2723
Db 1897 AGTCCCAAGTCAAAACCTGTGCAACAGAGCTGCAAGAGGGGAGGAGCCCTCCAA 1956
QY 2724 AATAACGGCAGAGAATAACAATGACAGAGAGTAACGGTGGAGCTTAAGTAGCCAAGGA 2783
Db 1957 AATAACGGCAGAGAATAACAATGACAGAGAGTAACGGTGGAGCTTAAGTAGCCAAGGA 2016
QY 2784 TTCTGAAACCTGGCATCCGTTCTGATGTCTGTGATGCTGATGCAATGATGCTGTCTGTG 2843
Db 2017 TTCTGAAACCTGGCATCCGTTCTGATGTCTGTGATGCTGATGCAATGATGCTGTCTGTG 2076
QY 2844 ACCCATCATATCCGCTACCAACCAATGCTTAATGCAATTTGGCAAGTTGATAGATATCT 2903
Db 2077 ACCCATCATATCCGCTACCAACCAATGCTTAATGCAATTTGGCAAGTTGATAGATATCT 2136
QY 2904 TTCCAAGATCGTTGTCTGTGACGTGGCCATGACATCATCAAGTCAATTAATTAATTTT 2963
Db 2137 TTCCAAGATCGTTGTCTGTGACGTGGCCATGACATCATCAAGTCAATTAATTAATTTT 2196
QY 2964 GGAATGAATCTCGATCATGCGCAGGAATTCATTAATCTAATCTGTGGAATTCGGCAGCCCAA 3023
Db 2197 GGAATGAATCTCGATCATGCGCAGGAATTCATTAATCTAATCTGTGGAATTCGGCAGCCCAA 2256
QY 3024 TACGGAGACAGAAAGTTTCATCATGTCATGCGGAGAGAAAGGATTAACCTTTGATA 3083
Db 2257 TACGGAGACAGAAAGTTTCATCATGTCATGCGGAGAGAAAGGATTAACCTTTGATA 2316
QY 3084 AATATCATGTACGCTTGGCCAAAGATGACCAACTCCCTCGAGGATTAACCAATGAA 3143
Db 2317 AATATCATGTACGCTTGGCCAAAGATGACCAACTCCCTCGAGGATTAACCAATGAA 2376
QY 3144 CGGTTTGGAAATTCCTGGGTGATGCTGTTGTTGAAATTTCTGACCAAGCGTCCAATTTGTACTAT 3203
Db 2377 CGGTTTGGAAATTCCTGGGTGATGCTGTTGTTGAAATTTCTGACCAAGCGTCCAATTTGTACTAT 2436
QY 3204 TTGTTTCTAGTCTGGAAGAGAGGATTAGCAACCTATCGGATCGCATTTGTTTCAAGAT 3263
Db 2437 TTGTTTCTAGTCTGGAAGAGAGGATTAGCAACCTATCGGATCGCATTTGTTTCAAGAT 2496
QY 3264 CAGCACTTGCATGCTAGCAAGAAACTTGAATCGATGATCCATTTATGTTATGTTCTAC 3323
Db 2497 CAGCACTTGCATGCTAGCAAGAAACTTGAATCGATGATCCATTTATGTTATGTTCTAC 2556
QY 3324 GGGCTCTGACCTTTGTAGAGAACTCGGACCTTCGACATGCAATGGCCAAATTTGTTGAAGCG 3383
Db 2557 GGGCTCTGACCTTTGTAGAGAACTCGGACCTTCGACATGCAATGGCCAAATTTGTTGAAGCG 2616
QY 3384 TTAATAGGAGCTGTTTACTTTGGAGGAAAGCGCTGGAGAGAGCCAAAGCAGTTATTTGACGC 3443
Db 2617 TTAATAGGAGCTGTTTACTTTGGAGGAAAGCGCTGGAGAGAGCCAAAGCAGTTATTTGACGC 2676
QY 3444 TTGCTCTTTAATGATCGGACCTCGCGAGTCTGGCTCAATTTATCTCTCCACCCACTC 3503
Db 2677 TTGCTCTTTAATGATCGGACCTCGCGAGTCTGGCTCAATTTATCTCTCCACCCACTC 2736
QY 3504 CAACTACAAGAGCAATACTGATGACCAACTTATTTGAAACTTCTCCAGTTCTTACAAAAA 3563
Db 2737 CAACTACAAGAGCAATACTGATGACCAACTTATTTGAAACTTCTCCGGTTCTACAAAAA 2796
QY 3564 CTTTACTGAGTTTGAAGAGCAATTTGGAGTAATTTTTTACTCATGTTCCGACTTCTGGCAAGG 3623
```

```
Db 2797 CTTACTGAGTTTGAAGAGCAATTTGGAGTAATTTTTTACTCATGTTTCACTTCTGGCAGG 2856
QY 3624 GCATTCAATTTGAGAACTGTGGGATTTAAACCATCTGACCCCTAGGCCCAATACAGAGATG 3683
Db 2857 GCATTCAATTTGAGAACTGTGGGATTTAAACCATCTGACCCCTAGGCCCAATACAGAGATG 2916
QY 3684 GAATTCCTAGGTGACTTCCATTAATGCAACTGGTGAAGCCACAGAGTACTTATTTCAATTC 3743
Db 2917 GAATTCCTAGGTGACTTCCATTAATGCAACTGGTGAAGCCACAGAGTACTTATTTCAATTC 2976
QY 3744 CCAGATCATCATGAGGACACTTAACTTTGTTGCGAAGCTCTTTGGTGAATAATAGAACT 3803
Db 2977 CCAGATCATCATGAGGACACTTAACTTTGTTGCGAAGCTCTTTGGTGAATAATAGAACT 3036
QY 3804 CAGGCCAAGGTAGCGGAGGAGCTGGGCATGCGAGGATGAGCCATAACCAACGACCAAGACC 3863
Db 3037 CAGGCCAAGGTAGCGGAGGAGCTGGGCATGCGAGGATGAGCCATAACCAACGACCAAGACC 3096
QY 3864 AAGAGCCCTGTGGCGCTTGCACCAAGACCTTTGGCGGACCTTTTGGAAATCATTTATGCA 3923
Db 3097 AAGAGCCCTGTGGCGCTTGCACCAAGACCTTTGGCGGACCTTTTGGAAATCATTTATGCA 3156
QY 3924 GCGCTGTACACTGATAAGGATTTGGAAATGTTTATCTATCTTTTCAATGATGTTCTCTTT 3983
Db 3157 GCGCTGTACACTGATAAGGATTTGGAAATGTTTATCTATCTTTTCAATGATGTTCTCTTT 3216
QY 3984 CCACGATTTGAAAGAAATTCATTTTGAATCAGGATTTGAATGACCCCAAAATCCCAAGCTTCAG 4043
Db 3217 CCACGATTTGAAAGAGTTTCATTTTGAACAGGATTTGAATGACCCCAAAATCCCAAGCTTCAG 3276
QY 4044 CAGTGTGCTGTGACACTTAGGACAGAAAGAAAGAGCCAGACATTCCTCTGTACAAGACT 4103
Db 3277 CAGTGTGCTGTGACACTTAGGACAGAAAGAAAGAGCCAGACATTCCTCTGTACAAGACT 3336
QY 4104 CTGCAGACAGTGGGCCCATCCCATGCGAACCTACACTGTGGCTGTTTATTTCAAGGGA 4163
Db 3337 CTGCAGACAGTGGGCCCATCCCATGCGAACCTACACTGTGGCTGTTTATTTCAAGGGA 3396
QY 4164 GAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTTACAGAAAGCGGAAATGGGAGCAGCAATG 4223
Db 3397 GAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTTACAGAAAGCGGAAATGGGAGCAGCAATG 3456
QY 4224 GATGCGCTTGAATAATAATATATTTTCCCGAGATGGCCCATCAGAAAGCGGTTTCATCGAACGG 4283
Db 3457 GATGCGCTTGAATAATAATATATTTTCCCGAGATGGCCCATCAGAAAGCGGTTTCATCGAACGG 3516
QY 4284 AAGTACAGACAGAGGTTTAAAGAAATGAGGTGGGAAAGAGAGCATCAAGAGAGAGGCCA 4343
Db 3517 AAGTACAGACAGAGGTTTAAAGAAATGAGGTGGGAAAGAGAGCATCAAGAGAGAGGCCA 3576
QY 4344 GATGAGACTGAAGACATCAAGAAATAA 4370
Db 3577 GATGAGACTGAAGACATCAAGAAATAA 3603
```

RESULT 11

AAH13924

ID AAH13924 standard; cDNA; 2598 BP.

XX AAH13924;

XX XX

DT 26-JUN-2001 (first entry)

XX XX

DE Human cDNA sequence SEQ ID NO:10948.

XX XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX XX

OS Homo sapiens.

XX XX

PN BP1074617-A2.

XX XX

PD 07-FEB-2001.

Db 1381 TTATTTGGAGCGTTGCTCTTTATGATCGGACCTCGCGAAGCTCGGCTCAATTATCCT 1440
Qy 3492 CTCACCCACTCCAACTACAGAGCCAAATACGTGATCGACAACCTTATTTGAACCTTCTCCA 3551
Db 1441 CTCACCCCACTCCAACTACAGAGCCAAATACGTGATCGACAACCTTATTTGAACCTTCTCCG 1500
Qy 3552 GTTCTACAAAACCTTACTGAGTTTGAAGAGCAATTTGGAGTAAATTTTACTCATGTTTCCA 3611
Db 1501 GTTCTACAAAACCTTACTGAGTTTGAAGAGCAATTTGGAGTAAATTTTACTCATGTTTCCA 1560
Qy 3612 CTTCTGCGAAGGCAATTCACATTTGAGAACTGTGGGATTTAAACCATCTGACCCCTAGGCCAC 3671
Db 1561 CTTCTGCGAAGGCAATTCACATTTGAGAACTGTGGGATTTAAACCATCTGACCCCTAGGCCAC 1620
Qy 3672 AATCAGAGATGAATTCCTAGTGATCTCCATTAATGCAACTGTGTAGCCACAGAGTACTTA 3731
Db 1621 AATCAGAGATGAATTCCTAGTGATCTCCATTAATGCAACTGTGTAGCCACAGAGTACTTA 1680
Qy 3732 TTTCAATTCATTTCCAGATCATCATGAGGACACTTAACCTTTGTTGGAGCTCTTTGGTG 3791
Db 1681 TTTCAATTCATTTCCAGATCATCATGAGGACACTTAACCTTTGTTGGAGCTCTTTGGTG 1740
Qy 3792 AATAATAGAACTCAGCCCAAGGTAGCGGAGGAGCTGGGATCGAGGATGACGCATAAACC 3851
Db 1741 AATAATAGAACTCAGCCCAAGGTAGCGGAGGAGCTGGGATCGAGGATGACGCATAAACC 1800
Qy 3852 AACGACAGACCAAGAGGCTGTGGCGCTTCGCAACAGACCTTGGCGGACCTTTTGGAA 3911
Db 1801 AACGACAGACCAAGAGGCTGTGGCGCTTCGCAACAGACCTTGGCGGACCTTTTGGAA 1860
Qy 3912 TCATTTATTTGACGCGCTGTACACTGATTAAGGATTTGGAAATGTTTCATCTTTTCATGAT 3971
Db 1861 TCATTTATTTGACGCGCTGTACACTGATTAAGGATTTGGAAATGTTTCATCTTTTCATGAT 1920
Qy 3972 GTCTGCTCTTTCCAGCATTTGAAGAAATTCATTTTGAATCAGGATTTGGAATGACCCCAA 4031
Db 1921 GTCTGCTCTTTCCAGCATTTGAAGAGTTCATTTTGAACAGGATTTGGAATGACCCCAA 1980
Qy 4032 TCCAGCTTCAGCAGTGTGCTTGAACACTTGAAGACAGAGGAAAGAGCCAGACATTCCT 4091
Db 1981 TCCAGCTTCAGCAGTGTGCTTGAACACTTGAAGACAGAGGAAAGAGCCAGACATTCCT 2040
Qy 4092 CTGTACAGACTCTGACAGACGTGGGCGCCATCCATGCCGACCTACACTGTGGCTGTT 4151
Db 2041 CTGTACAGACTCTGACAGACGTGGGCGCCATCCATGCCGACCTACACTGTGGCTGTT 2100
Qy 4152 TATTTTCAAGGAGAAAGATAGGCTCTGGGAAGGACCAAGTATTCAGCAAGCGGAAATG 4211
Db 2101 TATTTTCAAGGAGAAAGATAGGCTCTGGGAAGGACCAAGTATTCAGCAAGCGGAAATG 2160
Qy 4212 GGAGCAGCAATGGATGCGCTTGAATAATATAATTTTCCAGATGCCCATCAGAGCGG 4271
Db 2161 GGAGCAGCAATGGATGCGCTTGAATAATATAATTTTCCAGATGCCCATCAGAGCGG 2220
Qy 4272 TTCATCGAAGCAAGTACAGACAGAGTATAAGAAATGAGTGGGAAAGAGAGCATCAA 4331
Db 2221 TTCATCGAAGCAAGTACAGACAGAGTATAAGAAATGAGTGGGAAAGAGAGCATCAA 2280
Qy 4332 GAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAGGAGGCGCATGCAAGTGTGGA 4391
Db 2281 GAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAGGAGGCGCATGCAAGTGTGGA 2340
Qy 4392 GTATTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCTAGTTTCTCTGC 4451
Db 2341 GTATTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCTAGTTTCTCTGC 2400
Qy 4452 AGACAATGAACGAAGTGTGCTCAATGAATAAATAACAGAGTCAAAATCGTATTGTGTT 4511
Db 2401 AGACAATGAATGAAGTGTGCTCAATGAATAAATAACAGAGTCAAAATCGTATTGTGTT 2460
Qy 4512 TTAATGATCTGTTTTAGCTGGATGCTTTTATTACAAAGTATTAGATTTTCTTCTATT 4571
Db 2461 TTAATGATCTGTTTTAGCTGGATGCTTTTATTACAAAGTATTAGATTTTCTTCTATT 2520

Qy 4572 TAACGAAAACCTTGACTTTGGTGAATGTGCATTTACTTCTCTTTATTTTGTCTTTAAATA 4631
Db 2521 TAACGAAAACCTTGACTTTGGTGAATGTGCATTTACTTCTCTTTATTTTGTCTTTAAATA 2580
Qy 4632 ATAAAAATTCAGAAGCAT 4649
Db 2581 ATAAAAATTCAGAAGCAT 2598
RESULT 12
AEA20644
ID AEA20644 standard; cDNA; 1784 BP.
XX AEA20644;
AC AEA20644;
XX DT 11-AUG-2005 (first entry)
XX Novel human polynucleotide SEQ ID NO 1338.
XX vulnary; CNS-gen.; gene therapy; diagnostic; forensic; mapping;
KW DNA purification; protein purification; osteoarthritis; antiarthritic;
KW osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;
KW periodontal disease; antiinflammatory; mouth disease; burns; injury;
KW peripheral neuropathy; alzheimers disease; neuroprotective; nootropic;
KW degeneration; parkinsons disease; antiparkinsonian; neurological disease;
KW cerebrovascular ischemia; cerebroprotective; vasotropic;
KW cardiovascular disease; autoimmune disease; immunosuppressive;
KW immune disorder; viral infection; virucide; infection; cancer;
KW cytostatic; neoplasm; ss.
XX Homo sapiens.
OS
XX WO2005049806-A2.
PN
XX 02-JUN-2005.
PD
XX 11-MAR-2004; 2004WO-US007412.
PF
XX 14-MAR-2003; 2003US-00389559.
PR
XX (NUVE-) NUVELO INC.
PA
XX Tang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y;
PI Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R;
PI Wehrman T, Weng G, Boyle B;
PI WPI; 2005-417730/42.
DR
XX New polynucleotide encoding a polypeptide with biological activity,
PT useful for treating a disease or disorder, e.g. osteoarthritis, burns,
PT CNS and peripheral disease, stroke, autoimmune disorders, viral
PT infection, or cancer.
XX
PS Example 3; SEQ ID NO 1338; 500pp; English.
XX
CC The invention describes a new isolated polynucleotide (I) encoding a
CC polypeptide with biological activity comprising: a nucleotide sequence of
CC SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes
CC to the sequence of (i) under stringent hybridization conditions; or a
CC nucleotide sequence having greater than 99% sequence identity with the
CC sequence of (i). Also described are: a(n) (expression)vector comprising
CC (i); a host cell genetically engineered to comprise (i) operatively,
CC associated with a regulatory sequence that modulates expression of the
CC polynucleotide in the host cell; an isolated polypeptide comprising a
CC sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide
CC is: a polypeptide encoded by (i); or a polypeptide encoded by any one of SEQ
CC polynucleotide hybridizing under stringent conditions with any one of SEQ
CC ID NOS: 1-567; a composition comprising the polypeptide of (3) and a
CC carrier; an antibody directed against the polypeptide of (3); a method
CC for detecting (I) in a sample; a method for detecting the polypeptide of
CC (3) in a sample; a method for identifying a compound that binds to the
CC polypeptide of (3); a method of producing the polypeptide of (3); and a

CC collection of polynucleotides, where the collection comprising of at
CC least one of SEQ ID NOS: 1-567. (I) is a polynucleotide comprising any of
CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological
CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:
CC 568-1134. All sequences are fully defined in the specification. The
CC sequences and methods are useful in diagnostics, forensic, and gene
CC mapping, in identifying of mutations responsible for genetic disorders or
CC other traits, in assessing biodiversity, and for producing many other
CC types of data and products dependent on DNA and amino acid sequences. The
CC composition and method are useful for treating a disease or disorder,
CC e.g. osteoporosis, osteoarthritis, periodontal disease, burns, CNS and
CC peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,
CC autoimmune disorders, viral infection, or cancer. This sequence
XX represents a novel polynucleotide of the invention.

SQ Sequence 1784 BP; 536 A; 363 C; 405 G; 480 T; 0 U; 0 Other;

Query Match 36.4%; Score 1735.6; DB 14; Length 1784;

Best Local Similarity 98.9%; P-red. No. 0;

Matches 1747; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY	2901	ACTTTCCAGATCGTTGCTGTGTGAGCTGGCCATGCACTCATCCAAGTCATCATTTAAAT	2960
DB	14	ACCTTGTAATACATGTCGCGGCAGCTGGCCATGCACTCATCCAAGTCATCATTTAAAT	73
QY	2961	TTTGGAAATCAATCTCATGTCAGCAATTCATTAATCTAACTGTGCAATTCGGCAGCCC	3020
DB	74	TTTGGAAATCAATCTCATGTCAGCAATTCATTAATCTAACTGTGCAATTCGGCAGCCC	133
QY	3021	AAATACGAGACAGAAAGTTCAATCATGCAATCGGAAGAAAGGATTAACACCTTG	3080
DB	134	AAATACGAGACAGAAAGTTCAATCATGCAATCGGAAGAAAGGATTAACACCTTG	193
QY	3081	ATAATATATCATGTCAGCCCTGGCCAAAGATACCCAACTCCCTCGAGGATTAACCAAT	3140
DB	194	ATAATATATCATGTCAGCCCTGGCCAAAGATACCCAACTCCCTCGAGGATTAACCAAT	253
QY	3141	GAACGGTTCGAATTCCTGGGTGATGCTGTTTGAATTTCTGACCAAGCTCCATTTGTAC	3200
DB	254	GAACGGTTCGAATTCCTGGGTGATGCTGTTTGAATTTCTGACCAAGCTCCATTTGTAC	313
QY	3201	TATTTGTTTCTAGTCTGGAAGAGGAGGATAGCAACCTATCGGACTGCCATTTGTTAG	3260
DB	314	TATTTGTTTCTAGTCTGGAAGAGGAGGATAGCAACCTATCGGACTGCCATTTGTTAG	373
QY	3261	AATCAGCACTTGCATGTAGCAAAAGAACTTTGAATCGATGATGCTGTATGCT	3320
DB	374	AATCAGCACTTGCATGTAGCAAAAGAACTTTGAATCGATGATGCTGTATGCT	433
QY	3321	CACGGGCTGACCTTTGTAGAGATCGGACCTTCGACATGCAATGCCAATTTGTTGAA	3380
DB	434	CACGGGCTGACCTTTGTAGAGATCGGACCTTCGACATGCAATGCCAATTTGTTGAA	493
QY	3381	GCCTTAATAGGAGCTGTTTACTTGGAGGAAGCTTGGAGGAAGCCCAAGCAGTTATTGGA	3440
DB	494	GCCTTAATAGGAGCTGTTTACTTGGAGGAAGCTTGGAGGAAGCCCAAGCAGTTATTGGA	553
QY	3441	CGTTGCTCTTTAATGATCGGACCTCGCGGAAGTCTGGCTCAATTAATCTCTCCACCA	3500
DB	554	CGTTGCTCTTTAATGATCGGACCTCGCGGAAGTCTGGCTCAATTAATCTCTCCACCA	613
QY	3501	CTCCAACTACAGAGCCAAATCTGATCGCACTTTATTGAAACTTCTCCAGTTCTACAA	3560
DB	614	CTCCAACTACAGAGCCAAATCTGATCGCACTTTATTGAAACTTCTCCAGTTCTACAA	673
QY	3561	AAACTTACTGAGTTTGAAGAGCAATTTGGAGTAATTTTACTCATGTTTGGACTTTGGCA	3620
DB	674	AAACTTACTGAGTTTGAAGAGCAATTTGGAGTAATTTTACTCATGTTTGGACTTTGGCA	733
QY	3621	AGGGCAATTCATGTAGAACTGTGGGATTTAAACATCTGACCCCTAGGCCCAATTCAGAA	3680
DB	734	AGGGCAATTCATGTAGAACTGTGGGATTTAAACATCTGACCCCTAGGCCCAATTCAGAA	793

RESULT 13

ACF25385

ID ACF25385 standard; DNA; 1626 BP.

XX

QY	3681	ATGGAATTCCTAGGTGAGCTCCATAATGCAACTGGTAGCCACAGAGTACTTATTCATCAT	3740
DB	794	ATGGAATTCCTAGGTGAGCTCCATAATGCAACTGGTAGCCACAGAGTACTTATTCATCAT	853
QY	3741	TTCCAGATCATCATGAGGACACATTAATTTGTTGCGAAGCTCTTTGGTGAATATAGA	3800
DB	854	TTCCAGATCATCATGAGGACACATTAATTTGTTGCGAAGCTCTTTGGTGAATATAGA	913
QY	3801	ACTCAGGCCCAAGGTAGCGGAGGAGCTGGGCATGCGAGGAGTAGCCCATACCAACGACAAG	3860
DB	914	ACTCAGGCCCAAGGTAGCGGAGGAGCTGGGCATGCGAGGAGTAGCCCATACCAACGACAAG	973
QY	3861	ACCAAGAGCCCTGTGCGCTTCGCCAACAGACCTTTGGCGGACCTTTTGAATCATTTATT	3920
DB	974	ACCAAGAGCCCTGTGCGCTTCGCCAACAGACCTTTGGCGGACCTTTTGAATCATTTATT	1033
QY	3921	GCAGCGCTGTACACTGATTAAGGATTTGGNAATATGTTTACATCTTTCATGAATGCTGCTTC	3980
DB	1034	GCAGCGCTGTACACTGATTAAGGATTTGGNAATATGTTTACATCTTTCATGAATGCTGCTTC	1093
QY	3981	TTTCCACGATTCGAAAGAAATTCATTTTGAATCAGGATTTGGAATGATTTTACATCTTTCATGAATGCTGCTTC	4040
DB	1094	TTTCCACGATTCGAAAGAAATTCATTTTGAATCAGGATTTGGAATGATTTTACATCTTTCATGAATGCTGCTTC	1153
QY	4041	CAGCAGTGTGCTTGACACTTTAGGACAGAAAGAGCCAGACATTCCTCTGTACAAG	4100
DB	1154	CAGCAGTGTGCTTGACACTTTAGGACAGAAAGAGCCAGACATTCCTCTGTACAAG	1213
QY	4101	ACTCTGCAGACAGTGGGCCCATCCCATGCCCAACCTACACTGTGCTGTTTATTTCAAG	4160
DB	1214	ACTCTGCAGACAGTGGGCCCATCCCATGCCCAACCTACACTGTGCTGTTTATTTCAAG	1273
QY	4161	CGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTTACAGAAAGCGGAAATGGGAGCAGCA	4220
DB	1274	CGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTTACAGAAAGCGGAAATGGGAGCAGCA	1333
QY	4221	ATGGATGCGCTTTGAAAAATATTAATTTTCCAGATGCGCCCATCAGAAAGCGGTTTCATCGAA	4280
DB	1334	ATGGATGCGCTTTGAAAAATATTAATTTTCCAGATGCGCCCATCAGAAAGCGGTTTCATCGAA	1393
QY	4281	CGGAGTACAGACAGAGTTAAAGAAATGAGGTGGGAAAGAGGAGCATCAAGAGAGAGAG	4340
DB	1394	CGGAGTACAGACAGAGTTAAAGAAATGAGGTGGGAAAGAGGAGCATCAAGAGAGAGAG	1453
QY	4341	CCAGATGAGACTGAAAGACATCAAGAAATAAAGAGGCGCATGCAAGTGTGGAGTATTTACT	4400
DB	1454	CCAGATGAGACTGAAAGACATCAAGAAATAAAGAGGCGCATGCAAGTGTGGAGTATTTACT	1513
QY	4401	TGCTCAGTAACTGTGACTGTTCTATTGAGACCTAGCTAGCTAGTTTCTCCTGCAGACAATGA	4460
DB	1514	TGCTCAGTAACTGTGACTGTTCTATTGAGACCTAGCTAGTTTCTCCTGCAGACAATGA	1573
QY	4461	ACGAGTGTGCTCATTGAAATAAATACAGAGTCAATCGCTATGCTGTTTAAATGATC	4520
DB	1574	ACGAGTGTGCTCATTGAAATAAATACAGAGTCAATCGCTATGCTGTTTAAATGATC	1633
QY	4521	TGTTTTTACGTGATGCTTTTATTACAAAGTATTAGATTTTCTTCTATTAAACCGGAAA	4580
DB	1634	TGTTTTTACGTGATGCTTTTATTACAAAGTATTAGATTTTCTTCTATTAAACCGGAAA	1693
QY	4581	ACTTGACTTTGGTGAATGTCATTAATCTCTTTTATTTTGTCTTTTAAATATAAATTC	4640
DB	1694	ACTTGACTTTGGTGAATGTCATTAATCTCTTTTATTTTGTCTTTTAAATATAAATTC	1753
QY	4641	AGAAGCATATTTCTATGTCGAATAGA	4666
DB	1754	AGAAGCATATTTCTATGTCGAATAGA	1779

Db 1441 GGACCAAGTATTTCAGCAAGCGAAATGGAGCAGCAATGGATGCGCTTCAAAAATATAAT 1500
Qy 4245 TTTCCCCAGATGCCCATCAGAGCGGTTTCATCGAACGAGTACAGACAGAGTTAAA 4304
Db 1501 TTTCCCCAGATGCCCATCAGAGCGGTTTCATCGAACGAGTACAGACAGAGTTAAA 1560
Qy 4305 GAAATGAGTGGAAGAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAG 4364
Db 1561 GAAATGAGTGGAAGAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAG 1620
Qy 4365 AAATAA 4370
Db 1621 AAATAA 1626

RESULT 14
AAS33094
ID AAS33094 standard; cdNA; 1455 BP.

AC AAS33094;

DT 04-DEC-2001 (first entry)

XX DNA encoding human secreted protein, Seq ID No 53.

XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing; ss.

XX Homo sapiens.

XX WO200155326-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001347.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-022513P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.

PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	08-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-02559678P.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
XX	Rosen CA, Barash SC, Ruben SM;	
DR	WPI; 2001-451931/48.	
DR	P-PSDB; AAU20385.	
XX		
PT	New nucleic acids and polypeptides, useful for diagnosing, preventing or	
PT	treating medical conditions.	
XX		
PS	Claim 1; SEQ ID NO 53; 753pp; English.	
XX		
CC	The invention relates to novel isolated nucleic acid molecules (I)	
CC	encoding human secreted proteins (II). (I) and (II) are used to prevent,	
CC	treat or ameliorate a medical condition in e.g. humans, mice, rabbits, in	
CC	goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in	
CC	the prevention, treatment and diagnosis of diseases associated with	
CC	inappropriate expression of secreted proteins. (I) and complementary	
CC	sequences may also be used as DNA probes in diagnostic assays (e.g.	
CC	polymerase chain reactions (PCR)) to detect and quantitate the presence	
CC	of similar nucleic acid sequences in samples, and so which patients may	
CC	be in need of restorative therapy. (II) may also be used as antigens in	
CC	the production of antibodies and in assays to identify modulators	
CC	(agonists and antagonists) of the expression and activity of the secreted	
CC	proteins. The anti-(II) antibodies and antagonists may also be used to	
CC	down regulate expression and activity of (II). The anti-(II) antibodies	
CC	may also be used as diagnostic agents for detecting the presence of (II)	
CC	in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The	
CC	disorders include for example: immune/autoimmune diseases (e.g. HIV	
CC	(human immunodeficiency virus) infections, anaemia, rheumatoid arthritis	
CC	and multiple sclerosis), cancers and hyperproliferative disorders (e.g.	
CC	melanomas, neoplasms of the breast or liver, Sezary syndrome and	
CC	Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,	
CC	Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/	
CC	cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and	
CC	thrombosis), infections caused by bacteria, viruses and fungi and ocular	
CC	disorders (e.g. corneal infections). (I) and (II), agonists, antagonists	
CC	and antibodies can also be used to promote wound healing, maintain organs	
CC	before transplantation, and support cell culture of primary tissues.	

Query Match

30.2%; Score 1439.8; DB 4; Length 1455;

3

274

PROPERTY

COOPERATIVE

II

111111

[illegible]

Db 1022 ATGGCCCATCAGAACGGTTTCATCGAACGGAAGTACAGACAAGAGTTAAAAAGAAATGAGG 1081
QY 4314 TGGGAAGAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGG 4373
Db 1082 TGGGAAGAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGG 1141
QY 4374 AGGGCATCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACTGTTGTCTATTGAGAC 4433
Db 1142 AGGGCATCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACTGTTGTCTATTGAGAC 1201
QY 4434 CTAGCCTAGTTTCTTCGACAGACAATGAACGAAGTGTGCTCATTTGAATAAATAFACAGAT 4493
Db 1202 CTAGCCTAGTTTCTTCGACAGACAATGAACGAAGTGTGCTCATTTGAATAAATAFACAGAT 1261
QY 4494 CAAATCGCTATTGTTGTTTAAATGATCTGTTTGTAGCTGGATGCTTTTATTACAAAGTA 4553
Db 1262 CAAATCGCTATTGTTGTTTAAATGATCTGTTTGTAGCTGGATGCTTTTATTACAAAGTA 1321
QY 4554 TTAGATTTTCTTCTATTAAACGGAACCTTGACTTTGGTGAATGTGCATTACTTCCTTT 4613
Db 1322 TTAGATTTTCTTCTATTAAACGGAACCTTGACTTTGGTGAATGTGCATTACTTCCTTT 1381
QY 4614 TATTTTGTCTTTTAAATAATAAATAATCAAGAACATATTCTATGTGGAATAGATCCTGTT 4673
Db 1382 TATTTTGTCTTTTAAATAATAAATAATCAAGAACATATTCTATGTGGAATAGATCCTGTT 1441
QY 4674 TTTCATC 4681
Db 1442 TTTCAC 1449

RESULT 15
AAS34779
ID AAS34779 standard; cDNA; 1458 BP.
XX AAS34779;
AC AAS34779;
XX 04-DEC-2001 (first entry)
DE cDNA encoding novel human neoplastic disease associated polypeptide #13.
XX Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX Homo sapiens.
XX WO20015163-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001358.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 15-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225269P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0233968P.
PR 14-SEP-2000; 2000US-0233988P.
PR 14-SEP-2000; 2000US-0233999P.
PR 14-SEP-2000; 2000US-0234000P.
PR 14-SEP-2000; 2000US-0234010P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 08-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX P-PSDB; AAU21580.
XX WPI; 2001-465558/50.
XX P-PSDB; AAU21580.
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid arthritis.
XX
PS Claim 4; SEQ ID NO 23; 687pp; English.
XX
XX The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
CC sequences encoding for these polypeptides. The sequences of the invention
CC are useful in the diagnosis, treatment, prevention and/or prognosis of
CC disorders involving neoplastic disease such as hyperproliferative
CC disorders (e.g. leukemia, bone cancer, bladder cancer, brain stem
CC glioma, adult liver cancer, childhood cerebellar astrocytoma, or
CC Hodgkin's lymphoma). The sequences of the invention may also be useful
CC for treating other disorders such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, pulmonary disorders, cardiovascular disorders and renal
CC disorders. The polynucleotide sequences of the invention are also useful
CC in gene therapy. AAS34767-AAS35050 represent cDNA sequences encoding for
CC the novel human neoplastic disease associated polypeptides of the

CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1458 BP; 447 A; 291 C; 334 G; 386 T; 0 U; 0 Other;
Query Match 29.4%; Score 1398.6; DB 4; Length 1458;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1412; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 3234 GCAACCTATCGGACTGCCATTGTTGAGAATCAGCACCTTCCCATCTGCTAGCAAGAAACTT 3293
DB 24 GCAACCTATCGGACTGCCATTGTTGAGAATCAGCACCTTCCCATCTGCTAGCAAGAAACTT 83
QY 3294 GAACTGGATCCATTATGCTGTATGCTCAGCGGCTTGACCTTTGTAGAGAAATCGACCTT 3353
DB 84 GAACTGGATCCATTATGCTGTATGCTCAGCGGCTTGACCTTTGTAGAGAAATCGACCTT 143
QY 3354 CGACATGCAATGGCCAAATGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGGAGC 3413
DB 144 CGACATGCAATGGCCAAATGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGGAGC 203
QY 3414 CTGGAGGAAGCCAAAGCAGTTATTTGGACGCTTGCTCTTTAATGATCCGGACCTCGCGCAA 3473
DB 204 CTGGAGGAAGCCAAAGCAGTTATTTGGACGCTTGCTCTTTAATGATCCGGACCTCGCGCAA 263
QY 3474 GTCTGCTCAATPATCTCTCCACCACTCCAACTCAAGAGCAAAATCTGATCGACAA 3533
DB 264 GTCTGCTCAATPATCTCTCCACCACTCCAACTCAAGAGCAAAATCTGATCGACAA 323
QY 3534 CTTATTTGAACCTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAACTTGGAGTA 3593
DB 324 CTTATTTGAACCTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAACTTGGAGTA 383
QY 3594 ATTTTACTCATGTTCCGACTTCTGGCAAGGCAATTCACATTGAGAACTGTGGATTTAAC 3653
DB 384 ATTTTACTCATG-TGACTTCTGGCAAGGCAATTCACATTGAGAACTGTGGATTTAAC 442
QY 3654 CATCTGACCTAGGCCCAATCAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTG 3713
DB 443 CATCTGACCTAGGCCCAATCAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTG 502
QY 3714 GTAGCCACAGAGTACTTATTCATTCTTCCAGATCATCATGAGGACACTTAACTTTG 3773
DB 503 GTAGCCACAGAGTACTTATTCATTCTTCCAGATCATCATGAGGACACTTAACTTTG 562
QY 3774 TTGCGAAGCTCTTTGGTGAATTAATAGAACTCAGGCCAAGTACGCGAGGAGCTGGGCATG 3833
DB 563 TTGCGAAGCTCTTTGGTGAATTAATAGAACTCAGGCCAAGTACGCGAGGAGCTGGGCATG 622
QY 3834 CAGGAGTACGCCATTAACCAACGACAGCAAGCAAGAGGCTGTGGCGCTTCGACCAAGACC 3893
DB 623 CAGGAGTACGCCATTAACCAACGACAGCAAGCAAGAGGCTGTGGCGCTTCGACCAAGACC 682
QY 3894 TTGGCGGACCTTTTGAATCATTTATGAGCGCTGTACACTGATAGGATTTGGATAT 3953
DB 683 TTGGCGGACCTTTTGAATCATTTATGAGCGCTGTACACTGATAGGATTTGGATAT 742
QY 3954 GTTTCATCTTTCATGAATGCTGCTCTTTTCCAGGATTCAGAAATTCATTTTGAATCAG 4013
DB 743 GTTTCATCTTTCATGAATGCTGCTCTTTTCCAGGATTCAGAAATTCATTTTGAATCAG 802
QY 4014 GATTGGAATGACCCCAAAATCCAGCTTTCAGCAGTGTGCTTGAACCTTAGGACAGAAGGA 4073
DB 803 GATTGGAATGACCCCAAAATCCAGCTTTCAGCAGTGTGCTTGAACCTTAGGACAGAAGGA 862
QY 4074 AAAGAGCCAGACATTCCTCTGTACAGACTCTGCAGACAGTGGGCCATCCCATGCCCGA 4133
DB 863 AAAGAGCCAGACATTCCTCTGTACAGACTCTGCAGACAGTGGGCCATCCCATGCCCGA 922
QY 4134 ACCTACACGTGGCTGTTTATTTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGT 4193
DB 923 ACCTACACGTGGCTGTTTATTTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGT 982

QY	4194	ATTACGAAAGCGGAATGGAGCAGCAATGGATGCGCTTGAAAAATATAATTTTCCCCAG	4253
Db	983		
QY	4254	ATGGCCCATCAGAACGGTTTCATCGAAACGGAAGTACAGACAAGAGTTAAAAAGAAATGAGG	4313
Db	1043		
QY	4314	TGGGAAAGAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGG	4373
Db	1103		
QY	4374	AGGGCATGCAAGTGTGGAGTATTACTTGTCTAGTAACTGTGACTGTTGTCTATTGAGAC	4433
Db	1163		
QY	4434	CTAGCCTAGTTTTCTCGCAGACAAATGAACGAAGTGTGCTCATTTGAAATAAAATACAGAGT	4493
Db	1223		
QY	4494	CAAATCGCTATTGTGTGTTTAAATGATCTGTGTTTGTAGTGGATGGTCTTTATTACAAAGTA	4553
Db	1283		
QY	4554	TTAGATTTTCTTCTATTAAACGGAACAACTTGACTTTGGTGAATGTGCATTACTTCCTTT	4613
Db	1343		
QY	4614	TATTTTGTCTTTTAAATAATAAAATTCAGAAGCATA	4650
Db	1403		
		TATTTTGTCTTTTAAATAATAAAATTCAGAAGCATA	1439

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 14:31:04 ; Search time 17358 Seconds
(without alignments)
12840.976 Million cell updates/sec

Title: US-10-774-974-1
Perfect score: 4764
Sequence: 1 ctgtctgttgtaacctgcgcta.....aaaaaaaaaaaaaaaaaaaaa 4764

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1659	34.8	1920	11	DQ052678 Homo sapi
2	1520.4	31.9	3271	4	AK077549 Mus muscu
3	1496.2	31.4	1920	11	DQ052679 Pan trogl
4	1057.2	22.2	1443	10	AY411000 Homo sapi
5	994.8	20.9	1439	10	AY411001 Pan trogl
6	934.8	19.6	1094	7	CR753926 CR753926
7	899.2	18.9	1081	7	CR755186 CR755186
8	851.4	17.9	904	5	BQ883603 AGENCOURT
9	818.8	17.2	896	1	AU120566 AU120566
10	797.6	16.7	826	7	CR977645 CR977645
11	793.8	16.7	867	1	AU142521 AU142521
12	793.2	16.6	1436	10	AY411002 Mus muscu
13	763.2	16.0	826	6	CA454899 AGENCOURT
14	755.8	15.9	888	5	BUI55281 AGENCOURT
15	752.6	15.8	794	3	BI753613 603028687
16	741.2	15.6	783	2	BI255087 602975932
17	738	15.5	915	2	BE798610 601581671
18	732.2	15.4	739	8	DR423059 nav20h04
19	719.4	15.1	940	7	CR981843 CR981843
20	716.8	15.0	803	2	BI253637 602975314
21	714.8	15.0	718	6	CF129160 UI-HF-ES0
22	712.4	15.0	836	5	BU053284 UI-M-FC0

23	699.2	14.7	937	5	BQ891765	BQ891765 AGENCOURT
24	693.6	14.6	902	5	BUI62344	BUI62344 AGENCOURT
25	690	14.5	913	5	BQ900202	BQ900202 AGENCOURT
26	682	14.3	822	8	BX205797	BX205797 MNS09546
27	678.2	14.2	770	2	BG777150	BG777150 602664366
28	676.6	14.2	710	5	BUI619068	BUI619068 UI-H-FH1-
29	675.2	14.2	831	3	BP157462	BP157462 BP157462
30	668.2	14.0	960	5	BQ715944	BQ715944 AGENCOURT
31	667.6	14.0	945	6	CF585154	CF585154 AGENCOURT
32	659.4	13.8	889	6	CB192636	CB192636 AGENCOURT
33	658.6	13.8	934	2	BE792065	BE792065 601581771
34	656	13.8	669	1	AW246359	AW246359 2822248.5
35	654.8	13.7	728	2	BG031673	BG031673 602300826
36	654.4	13.7	794	7	CN531342	CN531342 UI-M-HQ0-
37	650.2	13.6	793	5	BU708701	BU708701 UI-M-F10-
38	646.8	13.6	875	6	CF108304	CF108304 Shultzomi
39	643.4	13.5	775	6	CF742155	CF742155 UI-M-HB0-
40	641.4	13.5	813	7	CN528777	CN528777 UI-M-HQ0-
41	637.8	13.4	713	7	CK833369	CK833369 4057248.8
42	637	13.4	787	2	BE545843	BE545843 601071744
43	633.6	13.3	662	6	CD702126	CD702126 EST18650
44	633.6	13.3	943	5	BQ959595	BQ959595 AGENCOURT
45	626.4	13.1	643	3	BI334592	BI334592 602999529

ALIGNMENTS

RESULT 1
DQ052678 1920 bp DNA linear GSS 02-JUN-2005
LOCUS Homo sapiens RNASE3L gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION DQ052678
VERSION DQ052678.1 GI:66898625
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (er) PLOS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1920)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source
1..1920
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
<1..>1920
/gene="RNASE3L"
/locus_tag="HC13224"

Query Match 34.8%; Score 1659; DB 11; Length 1920;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1768; Conservative 0; Mismatches 123; Indels 93; Gaps 1;

QY 273 ATGTGCTTCCACCCGGGACGAGGGCGTCCCGAGGACGAGGAGCAATGGAGCCAGACCC 332
Db 1 ATGTGCTTCCACCCGGGACGAGGGGTCTCCCGAGGACGAGGAGCAATGGAGCCAGACCC 60
QY 333 TCAGCACCAATCCCTTTAGGCCCCAAATCTGAGGCTGCTTCACTCAGCAGCCTCTGTG 392
Db 61 TCAGCACCAATCCCTTTAGGCCCCAAATCTGAGGCTGCTTCACTCAGCAGCCTCTGTG 120
QY 393 CAATATCAATATGAACCTCCAGTGCCTTCCACCACTTCTCAAACTCTCCAGCCCCC 452
Db 121 CAATATCAATATGAACCTCCAGTGCCTTCCACCACTTCTCAAACTCTCCAGCCCCC 180
QY 453 AATTTTCTCCCTCCAGCAGCACTTTGTACCCCTTCCCCCACTCCGTCAGCG 512
Db 181 AATTTTCTCCCTCCAGCAGCACTTTGTACCCCTTCCCCCACTCCGTCAGCG 240
QY 513 CAAGGCCCTCTTCCCCCTGCCCCAATCAGGCGGCTTTCCCAACCACTGAGGCGAC 572
Db 241 CAAGGCCCTCTTCCCCCTGCCCCAATCAGGCGGCTTTCCCAACCACTGAGGCGAN 300
QY 573 CCCTTCCAGTCTCTCTGTTTCTCCATGCCACCAATGCTTGTCTTAATAAC 632
Db 301 NNN 360
QY 633 CCCCCAGTCCCTGGGCGACCTCTGGACAAGGCACTTTCCCTTCACTGATGCCCTCCC 692
Db 361 NNN 420
QY 693 TCCATGCTCATCTCCCGCCCTCCAGTATGCGCGCAGAGTTAATTATCATGACCT 752
Db 421 TCCATGCTCATCTCCCGCCCTCCAGTATGCGCGCAGAGTTAATTATCATGACCT 480
QY 753 CCGGGCTATTCTACCACTCCAGCTCCAGCTCCAGTTTAAATAGTTTCCAGAACACCT 812
Db 481 CCGGGCTATTCTACCACTCCAGCTCCAGTTTAAATAGTTTCCAGAACACCT 540
QY 813 AGTTCTTCTGCGCCAGTCTAATAACAGAGTAGTCTCTATTTACAGACATCTCCCTCCA 872
Db 541 AGTTCTTCTGCGCCAGTCTAATAACAGAGTAGTCTCTATTTACAGACATCTCCCTCCA 600
QY 873 TACCACTCCCAAGGCTCCAGTGAGAGAAAGTCCCGAGAAAGCTGAAACATATGAT 932
Db 601 TACCACTCCCAAGGCTCCAGTGAGAGAAAGTCCCGAGAAAGCTGAAACATATGAT 660
QY 933 GACACAGGCAACGAGACCAAGTATGCGGCGAGGTGAGGATCCGTCCTGATCGG 992
Db 661 GACACAGGCAACGAGATCAAGTATGCGGCGAGGTGAGGATCCGTCCTGATCGG 720
QY 993 CGGAGCGAGGCGCGAGTCCGAGAGGAGAGACAGACAGCGCGTACAGATCTGATTAT 1052
Db 721 CGGAGCGAGGCGCGAGTCCGAGAGGAGAGACAGACAGCGCGTACAGATCTGATTAT 780
QY 1053 GACCGAGGAGAACACCACTCTCGCCACCGCAGCTACGAAACGAGCAGAGCGAGAACGG 1112
Db 781 GACCGAGGAGAACACCACTCTCGCCACCGCAGCTACGAAACGAGCAGAGCGAGAACGG 840
QY 1113 GAGAGACACAGGCAATCGAGAACACGAAAGATCACTCTCTGGAAGGTCTCAAAAAA 1172
Db 841 GAGAGACACAGGCAATCGAGAACACGAAAGATCACTCTCTGGAAGGTCTCAAAAAA 900
QY 1173 GAGTATAGAGATCTGGAAGGATACGGTTTATCGGTTTCTCTGAACCTGCTGGATGC 1232
Db 901 GAGTATAGAGATCTGGAAG----- 920
QY 1233 ACACCAAGATTACCTGGGAGATTATTAATAATACAGATTCTTGGGCCCCACCCCTGGAG 1292
Db 921 -----NNNNNN 927
QY 1293 ATTGTGAATCATCGTCCCAAGTAGGAGAGAGAGCTCGTTGGAGGAGAGAAAA 1352
Db 928 NNN 987

QY 1353 GACCGTTGAGTGACAAACAGAGTTCTGGCAAGACAAAGAACTATACCTCAATCAAGGAA 1412
Db 988 GACCGTTGAGTGACAAACAGAGTTCTGGCAAGACAAAGAACTATACCTCAATCAAGGAA 1047
QY 1413 AAAGAGCCCGAGGAGACCAATGCTGACAAAGAAATGAGGAGGAAAGAAAGAACTTCTTAAG 1472
Db 1048 AAAGAGCCCGAGGAGACCAATGCTGACAAAGAAATGAGGAGGAAAGAAAGAACTTCTTAAG 1107
QY 1473 CTTGTGTGATTCGATGCACTCATTCAGAAAATCTACTCTCCAGTGACCCCATGATCAG 1532
Db 1108 CTTGTGTGATTCGATGCACTCATTCAGAAAATCTACTCTCCAGTGACCCCATGATCAG 1167
QY 1533 GTGGGAGATTCTACAGTGTGTGAAACGAGTTCGTGACTTATATGACAAATTTGAG 1592
Db 1168 GTGGGAGATTCTACAGTGTGTGAAACGAGTTCGTGACTTATATGACAAATTTGAG 1227
QY 1593 GAGGAGTTGGGGAGCAGGCAAGAAAGGCTGCTCGGCTCCGTGGGAACTTCCA 1652
Db 1228 GAGGAGTTGGGGAGCAGGCAAGAAAGGCTGCTCGGCTCCGTGGGAACTTCCA 1287
QY 1653 AAGACAAAGCTCGATGAAGATTAGAGAGTTCCAGTGAATCCGAGTGTGATCTGATGAG 1712
Db 1288 AAGACAAAGCTCGATGAAGATTAGAGAGTTCCAGTGAATCCGAGTGTGATCTGATGAG 1347
QY 1713 GACAGCACCTGTTCTAGCAGCTCAGACTCTGAAGTTTGTGACGTATTTGACAGAAATCAAA 1772
Db 1348 GACAGCACCTGTTCTAGCAGCTCAGACTCTGAAGTTTGTGACGTATTTGACAGAAATCAAA 1407
QY 1773 CGCAAAAAGGCCCAACCTGACCGACTTTCATGATGAATTTGGTACAAACGATCCAGGCCAG 1832
Db 1408 CGCAAAAAGGCCCAACCTGACCGACTTTCATGATGAATTTGGTACAAACGATCCAGGCCAG 1467
QY 1833 ATCAATGATGGAACCACTCTGCAAAATGCGGCGAAAGCAAGACGACAGGAAATTTAGGCAC 1892
Db 1468 ATCAATGATGGAACCACTCTGCAAAATGCGGCGAAAGCAAGACGACAGGAAATTTAGGCAC 1527
QY 1893 AGCATTTATCTCGAGAGAGGCCATCAAGCCCTGCTGCTATGACCAACAAATGCTGGC 1952
Db 1528 AGCATTTATCTCGAGAGAGGCCATCAAGCCCTGCTGCTATGACCAACAAATGCTGGC 1587
QY 1953 AGACTTTTCCACTACCGGATCAAGTCTCCCGCCTACGAACTTTTAACTGACAGGCCA 2012
Db 1588 AGACTTTTCCACTACCGGATCAAGTCTCCCGCCTACGAACTTTTAACTGACAGGCCA 1647
QY 2013 ACTGTTATGAATACGATGATCAGAGTATATCTTTGAAAGGATTTCTATGTTTGCACAT 2072
Db 1648 ACTGTTATGAATACGATGATCAGAGTATATCTTTGAAAGGATTTCTATGTTTGCACAT 1707
QY 2073 GCGCCCTGACCAATATTCCACTGTTAAAGTAAATAGATTCAACATAGACTACAGATT 2132
Db 1708 GCGCCCTGACCAATATTCCACTGTTAAAGTAAATAGATTCAACATAGACTACAGATT 1767
QY 2133 CATTTCAATTGAAGAGATGATGCGCGAGAAATTTTGTGAAAGGGCTTGAATCTTTTCA 2192
Db 1768 CATTTCAATTGAAGAGATGATGCGCGAGAAATTTTGTGAAAGGGCTTGAATCTTTTCA 1827
QY 2193 CTGTTCTTATTCAGAGATATTTTGAATATATGATGCTGGAATCTTAAAGGCTTTGTTT 2252
Db 1828 CTGTTCTTATTCAGAGATATTTTGAATATATGATGCTGGAATCTTAAAGGCTTTGTTT 1887
QY 2253 GAAG 2256
Db 1888 CAGG 1891

RESULT 2
AK077549
LOCUS
DEFINITION
AK077549 3271 bp mRNA linear HTC 03-APR-2004
Mus musculus 8 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:5730445E20 product:weakly similar to
RIBONUCLEASE III (EC 3.1.26.3) (RNASE III) (P241) (Homo sapiens),
full insert sequence.
ACCESSION AK077549

```

VERSION
KEYWORDS
SOURCE
ORGANISM
MUS musculus
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Orawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3271)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozawa, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .3271

/misc_feature
1. .3271
/dev_stage="8 days embryo"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/tissue_type="whole body"
/db_xref="taxon:10090"
/db_xref="FANTOM_DB:5730445E20"
/mol_type="mRNA"
/strain="C57BL/6J"
/organism="Mus musculus"

ORIGIN
Query Match 31.9%; Score 1520.4; DB 4; Length 3271;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 1809; Conservative 0; Mismatches 316; Indels 19; Gaps 8;

Qy 1167 AAAAAAGATATAGAGATCTGGGAAGGAGTTACCGTTTATCGTTTGTCTCGAAGCTGCT 1226
Db 1 AAGAAAGAGTATAAGAGATCTGGGAAGGAGTTACCGTTTATCGTTTGTCTCGAAGCTGCT 60
Qy 1227 GGATGCACACAGAAATTACCTGGGGAGATTATTTAAATAACAGATTCTTGGGCCCCACCC 1286
Db 61 GGGTGCAACACAGAGTTGCCCTGGGGAGATGATTTAAACTACAGAGTCTTGGGCCCCGCC 120
Qy 1287 CTGGAGATTGTGAATCATCGCTCCCAAGTAGGAGGAAGAGAGAGTCTGTGGGAGGAA 1346
Db 121 CCGGAGAATGTGAATCATCGTTCTTCCAAGCAGGAGGAAGAGAGAGTCTGTGGGAGGAG 180
Qy 1347 GAAAAAGACCGTTGGAGTGACCAACAGAGTCTTCGCAAGACCAAGACATATACCTCAATC 1406
Db 181 GAAAAAGACAGATGGAGCGACAGCCAGGGCTCTGGCAAGAGAGAAATACACGTCCTATC 240
Qy 1407 AAGCAAAAAGA-GCCCGAGGAGACCATGCTCGTCAAGAAATCAGGAGGAAGAGAAAGACT 1465
Db 241 AAGAGAAAGAGGCGCAGAGAGGTGCTCCAGAGAGACGGAGGAGGAGGAGAGAGCT 300
Qy 1466 TCTTAAGCCTGTGTGGATTTCGATGCACTCAATTCAGAAAACTACTACTCCAGTGACCCCAT 1525
Db 301 CCTTAAGCCTGTGTGGATTTCGTTACACATTCGGAAGAGCTATTACTCCAGTGACCCCAT 360
Qy 1526 GGATCAGGTGGAGATCTTACAGTGGTTGGAAACAGTAGGCTTCGTGACTTATATGACAA 1585
Db 361 GGATCAGGTGGAGACTCGACTGTGTGAGGACAAAGCAGGCTCCGTGATTTGTATGACAA 420
Qy 1586 ATTTGAGGAGAGTTGGGGAGCAGGCAAGAAAGCCAAAGCTCTCGGCCCTCGTGGGA 1645
Db 421 ATTTGAGGAGAGATTTGGGGAGATAGGCAAGAGAGAGCCAAAGCTCCCGGCCCTCGTGGGA 480
Qy 1646 ACCTCCAAAGACGAAAGCTCGATGAAGATTTAGAGAGTTCCAGTGAATCCGAGTGTGAGTC 1705
Db 481 GCCTCCGAGACAAAGCTGGATGAAGATTTAGAGAGTTCCAGTGAATCCGAGTGTGAGAC 540
Qy 1706 TGATGAGGACGACACCTGTTCTAGCAGCTCAGACTCTGAAGTTTTTTCAGCTTATTCGAGA 1765
Db 541 TGACGACGACAGCAGCCTGTTTCGAGCAGCTCGGACTCGGAGGTGTTCGATGTTCATTCGAGA 600
Qy 1766 AATCAAGCGCAAAAGGCCCACTCGACCGACTTCATGATGAACCTTTGGTACAAAGATCC 1825
Db 601 GATTAAACGCAAAAGGCTCACCTCGACCGGCTTCATGATGAACCTTCGTTGTAACAAGACCC 660
Qy 1826 AGGCAGATGAATGATGGACCACTCT-GCAAAATGCAGCGCAAGAGGCAAGACGACAGGAA 1884
Db 661 AGGCAGATGACGATGACCGCTTTCGAAAAATGCAAGAGCCGACGACAGGAA 720
Qy 1885 TTAGGACACAGATTATTCCTGGAGAGAGGCGCATCAAGCCCTGTGCTCTATGACCAACA 1944
Db 721 TCGGCCACAGCATTTATCCCGGAGAGAGGCGCATCAAGCCCTGTGCTCTATGACCAACA 780
Qy 1945 ATGCTGGCAGACTTTTCCACTACCGGATCAGACTCTCCCGCCTACGACTTTTAACTG 2004
Db 1945 ATGCTGGCAGACTTTTCCACTACCGGATCAGACTCTCCCGCCTACGACTTTTAACTG 2004

```

Db 781 ACCTGGCCGGCTTTTCCATATCGGATCAACGCTCTCCCGCCTACCACTTCTTAACGTG 840

Qy 2005 ACAGGCCAACTGTTATAGAATACGATGATCAGGATATATCTTTGGAAGGATTTTCTATGT 2064

Db 841 ACAGGCCAACTGTTATAGAATATGATGATCAGGATATATTTTGAAGGATTTTCTATGT 900

Qy 2065 TTGCACATGCCCCCTGACCAATATTCACATGTGTAAAGTAATTAAGTTCACATAGACT 2124

Db 901 TTGCACATGCTCTCTGACCAATATTCACATGTGTAAAGTAATTCGATTCACATAGACT 960

Qy 2125 ACAGGATTCATTTCAATTTGAAGAGATGATGCGGAGATTTTGTGTAAGGCTTGAAC 2184

Db 961 ACAGGATTCATTTTATCGAGAGATGATGCTGAGGATTTTGTGTAAGGACTTGAAC 1020

Qy 2185 TCTTTTCACTGTTCTCTATTCAGAGATATTTTGGAAATATATGACTGGAATCTTAAAGGTC 2244

Db 1021 TGTTCATGTTCTCTATTCAGAGATATTTTGGAAATATATGACTGGAATCTTAAAGGTC 1080

Qy 2245 CTTTGTGTAAGACAGCCCTCTGCTGCTGCCAAGATTTCAATTCATGCCAGCTTTGTAA 2304

Db 1081 CTTTGTGTAAGACAGCCCTCTGCTGCTGCCAAGATTTCAATTCATGCCAGCTTTGTAA 1140

Qy 2305 GATTTCTTCAGATGAGGAGAAAGTGTCTGCTCCATGCAACAGATTTCTCTGACTTGT 2364

Db 1141 GATTTCTTCAGATGAGGAGAAAGTGTCTGCTCCATGCAACAGATTTCTCTGACTGCTGC 1200

Qy 2365 TAAGGTGAGCAAAAGCCCTGCTGCTGAGGAGATTTGCCAATATGCTTCAGTGGAGG 2424

Db 1201 TGCCTGAGCAAGGCTCTGCTGCTGCCAGGAGATTTGCCAATGCTTCAGTGGAGG 1260

Qy 2425 AGCTGGAGTGGCAGAAATATGACAGAAATGCAAGGATGATTTGTTACCAACCTGGGA 2484

Db 1261 AGCTGGAGTGGCAGAAATATGACAGGAGTGCAGGAGATTTGTTACCAACCTGGGA 1320

Qy 2485 CGAAA-CCAACTCTGCTCGTATCGA--TCACTGGATCGTGAACAGTTCAACCCCGA-- 2539

Db 1321 CGAAAACCGAGCTCTGCTCGCCATTAATCCAGCTGGATCGTGAACAGTTCAACCCCGAAA 1380

Qy 2540 -----TGTGATTACTTTTCGGATATCGTCCACTTTTGGGATAGCC--TGCACAGTT 2590

Db 1381 GTTAATCCCTTTTCCCAATATATCTTTTCCCTTTTGGATTTCCCTCCCGGCCAATTT 1440

Qy 2591 GAGTTATGAGGAGACCCACAGTACCAAAACTGTGGAAGATTTATGGAATCTTCGCCA 2650

Db 1441 AAGTAATCCAGAAACCCCATATCCAGAAATTTTGGAAAGTAAATTTGAAATTTGCCCC 1500

Qy 2651 CCTCTAGCAATAGTCCCAAGTCAACAACTGACAAACA--GAAGCTGGCAGAGG 2708

Db 1501 CTTCTTACCAACCTTTCCCAATCCAAATTAACAGCCAGAAAGTTGTTCCAGGG 1560

Qy 2709 GAGGAAGCCCTCCAAAAATACGGCAGAGAAATACAATGAGACGAGAAAGTAAACGGTGAG 2768

Db 1561 AGGAGGGCGTTCCAGAAATCCGCGCAGAGAACCCAAATCCGGCAGAAAGTTCCCGTTGGG 1620

Qy 2769 CTAAGTAGC-CAAGGATCTGGAATACTGGCATCCGTTCTGATGCTGTGACGATGCAAT 2827

Db 1621 TTAGTTAGCCCAAGGATTTTGGAAACTGGCAATTCGTTCTGATGCTGTGACGATGCAAT 1680

Qy 2828 GATGCTACTGTTCTGACCCATCATATCCGCTTACCACCAATGCCCTAAATGCAATTTGACAA 2887

Db 1681 GATGCTGCTGTTCTGACCCATCATATCCGGTACCATCAATGCGTAATGACCTGGACAA 1740

Qy 2888 GTTGATAGGATATCTTTTCCAAAGATGTTGTTCTGTTGACGCTGGCCATGACTCATCCAAAG 2947

Db 1741 GTTGATAGGATATCTTTTCCAAAGGCGTTGTTCTGCTCCAGCTTGGCATGACTCATCCGAG 1800

Qy 2948 TCATCATTTAAATTTTGGAAATGCTGATCATGCCAGGAATTCATTAATCTAACTGTGG 3007

Db 1801 TCACCATTTAAATTTTGGAAATGCTGATCATGCCAGGAATTTCTTTGTTCTAACTGTGG 1860

Qy 3008 AATTCGGACCCCAATACGGACAGAAAGTTTATCATGCAATGCAATGCGGAGAAAGG 3067

Db 1861 AATTCGACAAACCAATATGGAGACAGAAAGTTTATCATGCAATGCAATGCGGAGAAAGG 1920

Qy 3068 GATTAACACTTTGATATAATATCATGTACGCTTGGCCAGATGACCCAACTCCCTCGAG 3127

Db 1921 AATTAATACCTTAATAATAATATCATGTACGCTTGGCCAGATGATCCCACTCTTCAAG 1980

Qy 3128 GATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTGTTGAAATTTCTGACCAG 3187

Db 1981 GATTAATCAGCAAGAAAGTTGGAGTTCTGGGGATGCTGTTGTTGAGTTTCTGACCAG 2040

Qy 3188 CGTCCATTTGTACTATTGTTTCTCTAGTCTGGAAGAAGGAGGATTAGCACTATCGGAC 3247

Db 2041 TGTCCACCTGTACTACTGTTTCTTAGCTTGGAGGAAGGGGCTTGGCGACCTATCGGAC 2100

Qy 3248 TGCATTTGTCAGAAATCAGCACTTGCCTAGCTAGCAAAAGAAAC 3291

Db 2101 CGCATTTGTTTCAAAATCAGCACTTGCCTATGCTTGCAGAGGTAC 2144

RESULT 3

LOCUS DQ052679 1920 bp DNA linear GSS 02-JUN-2005

DEFINITION Pan troglodytes RNASE3L gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION DQ052679

VERSION DQ052679.1 GI:66898626

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 1920) Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civallo, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees (er) PLOS Biol. 3 (6), E170 (2005)

AUTHORS

TITLE

JOURNAL PUBMED 15869325

REFERENCE 2 (bases 1 to 1920) Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civallo, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. Direct Submission

AUTHORS

TITLE

JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES

source 1..1920

gene /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>1920

ORIGIN

1 /gene="RNASE3L" /locus_tag="HC13224"

Query Match 31.4%; Score 1496.2; DB 11; Length 1920;

Best Local Similarity 80.7%; Pred. No. 0;

Matches 1610; Conservative 0; Mismatches 292; Indels 93; Gaps 1;

Qy 273 ATGTGTTTCCACCCGGGACGAGGGGTCTCCCGAGACGAGGACATGGAGCCAGACCC 332

Db 1 ATGTGTTTCCACCCGGGACGAGGGGTCTCCCGAGACGAGGACATGGAGCCAGACCC 60

Qy 333 TCAGCACCATCTTTTAGGCCCCCAAAATCTGAGGCTTTCACCTTCAGCAGCCTCTGTG 392

Db 61 TCAGCACCATCTTTTAGGCCCCCAAAATCTGAGGCTTTCACCTTCAGCAGCCTCTGTG 120

Qy 393 CAATATCAATATGAACCTTCCAAGTCCCTTCCACACTTTTCTCAAACTTCCAGCCCC 452

Db 121 CAATATCAATATGAACCTCAAGTCCCTCCACCACTTTCTCGAACTCTCCAGCCCC 180
Qy 453 AATTTTCTCCCTCCACGACGAGCTTTGTATCCCTTCCCTCCACCGAGCTCTCGTCAGCG 512
Db 181 AATTTTCTCCCTCCACGACGAGCTTTGTATCCCTTCCCTCCACCGAGCTCTCGTCAGCG 240
Qy 513 CAAGGCTCTCTTCCCTCCCTCCCAATCAGGCGCTTTTCCCAACCAACGAGTCAAGGCAC 572
Db 241 CAAGGCTCTCTTCCCTCCCTCCCAATCAGGCGCTTTTCCCAACCAACGAGTCAAGGCAN 300
Qy 573 CCCTTCCAGTTCCTCTCTGTTTCTCCCTCCATGCCACCAACCAATGCTTGTCTTAATAAC 632
Db 301 NNN 360
Qy 633 CCCCCAGTCCCTGGGGCAGCTCTCTGGACAAGGCACTTCCCTTCATGATGCCCTCC 692
Db 361 NNN 420
Qy 693 TCCATGCTCATCCCCCGCCCTCCAGTCAATGCCGAGCAGGTTAATTATCAGTACCT 752
Db 421 NNN 480
Qy 753 CCGGCTATTCTCACCACTTCCCTCCCTCCAGTCTTAACTGTTTCCAGAAACAACT 812
Db 481 NNN 540
Qy 813 AGTTCTTTCTGCGCCAGTCTAATAACAGCAGTATGCTCTCAATTCAGACATCTCCCTCCA 872
Db 541 AGTTCTTTCTGCGCCAGTCTAATAACAGCAGTATGCTCTCAATTCAGACATCTCCCTCCA 600
Qy 873 TACCACCTCCCAAGGCTCCAGTGGAGAGGCTCCCAAGAGGCTGAACACTATGAT 932
Db 601 TACCACTCCCAAGGCTCCAGTGGAGAGGCTCCCAAGAGGCTGAACACTATGAT 660
Qy 933 GACCACAGGACCCAGACACAGTCTGCGGAGGTGAGAGGCTCCCTCGATCGG 992
Db 661 GACCACAGGACCCAGACACAGTCTGCGGAGGTGAGAGGCTCCCTCGATCGG 720
Qy 993 CGGAGCGAGCGCGAGTCCGACAGGAGAACAGACAGCGGTTACAGATCTGATTAT 1052
Db 721 CGGAGCGAGCGCGAGTCCGACAGGAGAACAGACAGCGGTTACAGATCTGATTAT 780
Qy 1053 GACGAGGAGACACACATCTCGCCACCGCAGTCTACGAAACGAGCAGAGCGAGACGG 1112
Db 781 GACGAGGAGACACACATCTCGCCACCGCAGTCTACGAAACGAGCAGAGCGAGACGG 840
Qy 1113 GAGAGACACAGGATCAGACACAGGAGTCAATCTCTGGAAGGTCTCAACAAAA 1172
Db 841 GAGAGACACAGGATCAGACACAGGAGTCAATCTCTGGAAGGTCTCAACAAAA 900
Qy 1173 GAGTATAAGAGATCTGGAAGGAGTTACGTTTATCGGTTTCTCTGAACTCTGCTGATGC 1232
Db 901 GAGTATAAGAGATCTGGAAG----- 920
Qy 1233 ACACGAGATTACCTGGGAGATTATTAATAATACAGATTCTTGGGGCCCAACCTGGAG 1292
Db 921 -----NNNNNN 927
Qy 1293 ATTGTGAATCATCGTCCCAAGTAGGAGAGAGAGCTCGTTGGGAGGAGAAAA 1352
Db 928 NNN 987
Qy 1353 GACCGTTGGAGTCAACACAGAGTTCTGGCAAGAACATATACCTCAATCAAGGAA 1412
Db 988 GACCGTGGAGTCAACACAGAGTTCTGGCAAGAACATATACCTCAATCAAGGAA 1047
Qy 1413 AAGAGCCCGAGGAGACCATGCTGACAGAAATGAGGAGGAGAGAACTCTTTAAG 1472
Db 1048 AAGAGCCCGAGGAGACCATGCTGACAGAAATGAGGAGGAGAGAACTCTTTAAG 1107
Qy 1473 CTTGTGTGAGTTCGATGCACTTTCAGAAACTTACTTCCAGTGAACCCATGATCAG 1532
Db 1108 CTTGTGTGAGTTCGATGCACTTTCAGAAACTTACTTCCAGTGAACCCATGATCAG 1167

Qy 1533 GTGGGAGATTCTACAGTGGTTGGAAACGAGTAGGCTTCGTGACTTATATGACAAAATTGAG 1592
Db 1168 GTGGGAGATTCTACAGTGGTTGGAAACGAGTAGGCTTCGTGACTTATATGACAAAATTGAG 1227
Qy 1593 GAGGAGTTGGGGAGCAGCAAGAAAGCCAAAGCTGCTCGGCTCCGTGGGAACCTCCA 1652
Db 1228 GAGGAGTTGGGGAGCAGCAAGAAAGCCAAAGCTGCTCGGCTCCGTGGGAACCTCCA 1287
Qy 1653 AAGACGAAGCTCGATGAAGATTTAGAGAGTTCCAGTGAATCCGAGTGTGAGTCTGATGAG 1712
Db 1288 AAGACGAAGCTCGATGAAGATTTAGAGAGTTCCAGTGAATCCGAGTGTGAGTCTGATGAG 1347
Qy 1713 GACAGCACCTGTTCTACAGCTCAGACTCTGAAGTTTTTGAAGTTTTTGAAGAAATCAAA 1772
Db 1348 GACAGNCCCTGTTCTAGCAGCTCAGACTCTGAAGTTTTTGAAGTTTTTGAAGAAATCAAA 1407
Qy 1773 CGCAAAAGCCCAACCTCGACCTGACCTTTCATGATGAATCTTGGTACAACTTCCAGGCCAG 1832
Db 1408 CGCAAAAGCCCAACCTCGACCTTTCATGATGAATCTTGGTACAACTTCCAGGCCAG 1467
Qy 1833 ATGAATGATGACCACTCTGCAAAATGAGCGCAAGGCAAGACGACAGGAATTAGGCAC 1892
Db 1468 ATGAATGATGACCACTCTGCAAAATGAGCGCAAGGCAAGACGACAGGAATTAGGCAC 1527
Qy 1893 AGCAATTTATCTGAGAGAGGCGCATCAAGCCCTGCTGCTCTATGACCAACAATGCTGGC 1952
Db 1528 AGCAATTTATCTGAGAGAGGCGCATCAAGCCCTGCTGCTCTATGACCAACAATGCTGGC 1587
Qy 1953 AGACTTTTCCACTACCGATCAGACTCTCCCGCTACGAACTTTTAACTGACAGGCCA 2012
Db 1588 AGACTTTTCCACTACCGATCAGACTCTCCCGCTACGAACTTTTAACTGACAGGCCA 1647
Qy 2013 ACTGTTATAGAATACGATGATCAGAGTATCTTTGAAGGATTTCTATGTTTGCACAT 2072
Db 1648 ACTGTTATAGAATACGATGATCAGAGTATCTTTGAAGGATTTCTATGTTTGCACAT 1707
Qy 2073 GCGCCCTGACCAATATTCACCTGTGTAAAGTAAATAGATTCAACATAGACTACAGATT 2132
Db 1708 GCGCCCTGACCAATATTCACCTGTGTAAAGTAAATAGATTCAACATAGACTACAGATT 1767
Qy 2133 CATTTCAATGAAGAGATGATGCGGAGAAATTTTGTGTAAAGGCTTGAACCTTTTCA 2192
Db 1768 CATTTCAATGAAGAGATGATGCGGAGAAATTTTGTGTAAAGGCTTGAACCTTTTCA 1827
Qy 2193 CTGTTCTTATTCAGAGATATTTTGGAAATATATGACTTGAAGCTCTTAAAGCTCTTTGTT 2252
Db 1828 CTGTTCTTATTCAGAGATATTTTGGAAATATATGACTTGAAGCTCTTAAAGCTTAACTATCTT 1887
Qy 2253 GAAGACAGCCCTCCC 2267
Db 1888 CAGGGNNNACTTTCC 1902

RESULT 4

AY411000
LOCUS
DEFINITION
Homo sapiens HCM4078 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY411000
VERSION
AY411000.1 GI:39766968
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 1443)
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1443)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..1443
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1443
/locus_tag="HCM4078"
ORIGIN
Query Match 22.2%; Score 1057.2; DB 10; Length 1443;
Best local Similarity 73.4%; Pred. No. 2e-262;
Matches 1059; Conservative 0; Mismatches 384; Indels 0; Gaps 0;
QY 2928 CTGCCCATGACTCATCCAAAGTCATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGCAGG 2987
DB 1 CTGCCCATGACTCATCCAAAGTCATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGCAGG 60
QY 2988 AATTCATTATCTAATCTGTGGAATTCGGCAGCCCAATACCGAGACAGAAAAGTTTCATCAC 3047
DB 61 AATTCATTATCTAATCTGTGGAATTCGGCAGCCCAATACCGAGACAGAAAAGTTTCATCAC 120
QY 3048 ATGCACATGCGGAGAAAGGAATTAACACCTTGATAAATATCATGTCAACGCTTGGCCAA 3107
DB 121 ATGCACATGCGGAGAAAGGAATTAACACCTTGATAAATATCATGTCAACGCTTGGCCAA 180
QY 3108 GATGACCAACTCCCTCGAGGATTAACCACAATGAACGGTTGGAAATTCCTGGGTGATGCT 3167
DB 181 NNN 240
QY 3168 GTTGTGAAATTCGACACAGCGTCCAATTTGACTATTTGTTTCTAGTCTGGAAGAAGGA 3227
DB 241 GTTGTGAAATTCGACACAGNN 300
QY 3228 GGATTAGCAACTATCGACTGCCATTGTTGAGAAATCAGACCTTGCCATGCTAGCAAAG 3287
DB 301 NNN 360
QY 3288 AAACCTTGAACTGGATCCCAATTTATGCTGATGCTCAACGGGCTGACCTTTGTAGAGAATCG 3347
DB 361 AAACCTTGAACTGGATCGAATTTATGCTGATGCTCAACGGGCTGACCTTTGTAGAGAATCG 420
QY 3348 GACCTTCGACATGCAATGCCCAATTTGAGCGTTAATAGGAGCTGTTACTTGAG 3407
DB 421 GACCTTCGACATGCAATGCCCAATTTGAGCGTTAATAGGAGCTGTTACTTGAG 480
QY 3408 GGAACGCTGGAGGAGCAACGAGTTATTTGACGCTTGTCTTTAATGATCCGAGACTG 3467
DB 481 NNN 540
QY 3468 CGCGAAGTCTGGCTCAATTTATCTCTCCACCCACTCCAACTACAAGAGCCAAATCTGAT 3527
DB 541 CGCGAAGTCTGGCTCAATTTATCTCTCCACCCACTCCAACTACAAGAGCCAAATCTGAT 600
QY 3528 CGACAACTTATTAAGAACTCTCCAGTCTCAAAAACCTTACTGAGTTGGAAGCAATTT 3587
DB 601 CGACAACTTATTAAGAACTCTCCAGTCTCAAAAACCTTACTGAGTTGGAAGCAATTT 660
QY 3588 GGAGTAATTTTACTCATGTTTCGACTTCTGGCAAGGGCAATTCACATTTGAGAACTGTGGGA 3647
DB 661 GGAGTAATTTTACTCATGTTTCGACTTCTGGCAAGGGCAATTCACATTTGAGAACTGTGGGA 720

QY 3648 TTTAACCATCTGACCCCTAGGCCACAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATG 3707
DB 721 TTTAACCATCTGACCCCTAGGCCACAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATG 780
QY 3708 CAACTGGTAGCACAAGAGTACTTATTCATTTCATTTCCAGATCATCATGAAGACACTTTA 3767
DB 781 CAACTGGTAGCACAAGAGTACTTATTCATTTCATTTCCAGATCATCATGAAGACACTTTA 840
QY 3768 ACTTTGTTGCGAAGCTCTTTGGTGAATAATAAGAACTCAGGCCAAGGTACGGGAGAGCTG 3827
DB 841 ACTNN 900
QY 3828 GGCATCGCAGGATACGCCATAACCAACGACAAGACGAGCGCTGTGGCGCTTCGCACC 3887
DB 901 NNN 960
QY 3888 AAGACCTTGGCGGACCTTTTGGAAATCATTTATTCAGCGCTGTACACTCATAGGATTTG 3947
DB 961 NNN 1020
QY 3948 GAATATGTTTCATACTTTTCATGAATGTCGCTCTTCCACGATTTGAAAGAAATTCATTTTG 4007
DB 1021 GAATATGTTTCATACTTTTCATGAATGTCGCTCTTCCACGATTTGAAAGATTTTCATTTTG 1080
QY 4008 AATCAGATTTGGAATGACCCCAATCCAGCTTCACAGTGTTCAGACATTTGACACTTAGGACA 4067
DB 1081 AATCAGATTTGGAATGACCCCAATCCAGCTTCACAGTGTTCAGACATTTGACACTTAGGACA 1140
QY 4068 GAAAGAAAAGAGCAGACATTCCTCTGTACAGACTCTGCAGACAGTGGCCCCCATCCCAT 4127
DB 1141 GAAGGAAAAGAGCAGACATTCCTCTGTACAGACTCTGCAGACAGTGGCCCCCATCCCAT 1200
QY 4128 GCCCGAACCTTACACTCTGGCTGTTTATTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGGA 4187
DB 1201 GCCCGAACCTTACACTCTGGCTGTTTATTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGGA 1260
QY 4188 CCAAGTATTTCAGCAAGCGGAAATGGGAGCAGCAATGCGTGTGAAAAATATAATTTT 4247
DB 1261 CCAAGTATTTCAGCAAGCGGAAATGGGAGCAGCAATGCGTGTGAAAAATATAATTTT 1320
QY 4248 CCCAGATGGCCCATCAGAAAGCGGTTTCATCGAAGCGGAGTACAGACAGAGTTAAAGAA 4307
DB 1321 CCCAGATGGCCCATCAGAAAGCGGTTTCATCGAAGCGGAGTACAGACAGAGTTAAAGAA 1380
QY 4308 ATGAGTGGGAAAGAGAGCATCAAGAGAGAGCCAGTACAGACTGAAAGACATCAAGAAA 4367
DB 1381 ATGAGTGGGAAAGAGAGCATCAAGAGAGAGCCAGTACAGACTGAAAGACATCAAGAAA 1440
QY 4368 TAA 4370
DB 1441 TAA 1443
RESULT 5
LOCUS AY411001
DEFINITION Pan troglodytes HCM4078 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411001
VERSION AY411001.1 GI:39766969
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarctomylres; Primates; Catarrhini;
Hominidae; Pan.
REFERENCE 1 (bases 1 to 1439)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 1439)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D., and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 source
 1..1439
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1439
 /locus_tag="HCM4078"
 ORIGIN
 Query Match 20.9%; Score 994.8; DB 10; Length 1439;
 Best Local Similarity 69.4%; Pred. No. 3.2e-246;
 Matches 999; Conservative 0; Mismatches 440; Indels 0; Gaps 0;
 2932 CCATGACTCATCAAGTCATCAATTTAAATTTTGGAAATGAATCCTGATCATGCCAGGAATT 2991
 Db 1 CCATGACTCATCAAGTCATCAATTTAAATTTTGGAAATGAATCCTGATCATGCCAGGAATT 60
 2992 CATTATCTAACTGTGGATTTGGGAGCCCAATACGGAGACAGAAAGTTTCATCATATGC 3051
 Db 61 CGTTGTCTAACTGTGGATTTGGGAGCCCAATACGGAGACAGAAAGTTTCATCATATGC 120
 3052 ACATCGGAAGAAAGGGAATTAACACCTTGATAATATCATGTACGCTTGCCCAAGATG 3111
 Db 121 ACATCGGAAGAAAGGNN 180
 3112 ACCCAACTCCCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTG 3171
 Db 181 NNNNNNNNNNNNNNGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTG 240
 3172 TTGAATTTCTGACGAGCTGCATTTGTACTATTGTTCTTCTAGTCTGGAAGAGGAT 3231
 Db 241 TTGAATTTCTAACGAGNN 300
 3232 TAGCAACTCTCGGACTGCATTGTTTCAGAAATCAGCACTTGCCATGCTAGCAAGAAAC 3291
 Db 301 NNN 360
 3292 TTGAATCGATCCATTTATGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACC 3351
 Db 361 TTGAATCGATCGATNNNTGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACC 420
 3352 TTGACATGCAATGCCCATTGTTTGAAGCGTTTAATAGGAGCTGTTTACTTTGGAGGAA 3411
 Db 421 TTGACATGCGNTGCGNNNTGTTTGAAGCAATTAATNNNNNNNNNNNNNNNNNNNNNN 480
 3412 GCCTGGAGGAGCCAGCAGTATTGTCAGCTGCTCTTAAATGATCCGGACCTGCGG 3471
 Db 481 NNN 540
 3472 AAGTCTGGCTCAATATATCTCTCCACCACTCAAGAGCAAAATCTGATCGAC 3531
 Db 541 AAGTCTGGCTCAATATATCTCTCCACCACTCAAGAGCAAAATCTGATCNAN 600
 3532 AACTTATTGAACCTTCTCCAGTTCTTCAAAAATTTACTGAGTTGAAGCAATTTGGAG 3591
 Db 601 NNCCTATTGAACCTTCTCCGGTTCTACNNAATTTACTGAGTTGAAGCAATTTGGAG 660
 3592 TAAATTTTACTCATGTTTCGACTTGGCAAGGGCAATTCACATTGGAATTTGAGATTTA 3651
 Db 661 TAAATTTTACTCATGTTTCGACTTGGCAAGGGCAATTCACATTGGAATTTGAGATTTA 720

QY 3652 ACCATCTGACCTTAGGCCCAATTCAGAGAAATGGAATTCCTAGTGAATCCATAATGCAAC 3711
 Db 721 ACCATCTGACCTTAGGCCCAATTCAGAGAAATGGAATTCCTAGTGAATCCATAATGCAAC 780
 QY 3712 TGGTAGCCACAGTACTTATTTCATTTCCTCCAGATCATCATGAAGACACTTAACTT 3771
 Db 781 TGGTAGCCACAGTACTTATTTCATTTCCTCCAGATCATCATGAAGACACTTAACTN 840
 QY 3772 TGTTCGGAAGCTCTTTTGGTGAATAATAGAACTCAGCCCAAGGTAGCGAGGAGCTGGCA 3831
 Db 841 NNN 900
 QY 3832 TGCAGAGTACGCCATAACCAAGCAAGACCAAGAGGCTGTGGCGCTTCGCAACAAGA 3891
 Db 901 NNN 960
 QY 3892 CCTTGGCGGACCTTTTGGAAATCAATTTATTCAGCGCTGTACACTGATAAGGATTTCGAAT 3951
 Db 961 NNN 1020
 QY 3952 ATGTTTCATCTTCATGAATGTCTCTCTTCCAGATTTGAAAGAAATTCATTTTGAATC 4011
 Db 1021 ATGTTTCATCTTCATGAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080
 QY 4012 AGGATTGGAATGACCCCAAAATCCAGCTTCAGAGCTGTGCTTGACACTTAGACAGAG 4071
 Db 1081 AGGATTGGAATGACCCCAAAATCCAGCTTCAGAGCTGTGCTTGACACTTAGACAGAG 1140
 QY 4072 GAAAAGCCAGACATCTCTGTACAGACTCTCCAGACACTGCGGCCCATCCCATGCC 4131
 Db 1141 GAAAAGCCAGACATCTCTGTACAGACTCTCCAGACACTGCGGCCCATCCCATGCC 1200
 QY 4132 GAACCTACACTGTGCTGTTTATTTCAAGGGAGAAAGAAATAGGCTGTGGAAAGACCAA 4191
 Db 1201 GAACCTACACTGTGCTGTTTATTTCAAGGGAGAAAGAAATAGGCTGTGGAAAGACCAA 1260
 QY 4192 GTATTTCAGCAACGGGAAATGGAGCAGCAATGGATGCGCTTGAAAAATATAATTTTCCC 4251
 Db 1261 GTATTTCAGCAACGGGAAATGGAGCAGCAATGGATGCGCTTGAAAAATATAATTTTCCC 1320
 QY 4252 AGATGGCCCATCAGAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTAAAGAAATGA 4311
 Db 1321 AGATGGCCCATCAGAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTAAAGAAATGA 1380
 QY 4312 GGTGGGAAAGAGACATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAA 4370
 Db 1381 GGTGGGAAAGAGACATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAA 1439
 RESULT 6
 CR753926
 LOCUS
 DEFINITION
 CR753926 Rattus norvegicus muscle Sprague-Dawley Rattus norvegicus
 cDNA clone GP00AA152A05, mRNA sequence.
 ACCESSION
 CR753926
 VERSION
 CR753926.1 GI:51865883
 KEYWORDS
 EST.
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.
 1 (bases 1 to 1094)
 CROSS, N., Tkatchenko, A.V., Pisani, D.F., Leclerc, L., Leger, J.J.,
 Marini, J.F. and Dechesne, C.A.
 Analysis of altered gene expression in rat soleus muscle atrophied
 by disuse
 J. Cell. Biochem. 83 (3), 508-519 (2001)
 11596118
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Campbell AP, Chenchik A, Mogadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Svetlov ED, Siebert PD. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A*. 1996; 93: 6025-30. Rats were female Sprague Dawley between 200 and 220g. Soleus muscle atrophy was performed by 14 days of hindlimb suspension."

ORIGIN

Query Match 18.9%; Score 899.2; DB 7; Length 1081;
Best Local Similarity 90.7%; Pred. No. 2e-221;
Matches 980; Conservative 0; Mismatches 98; Indels 2; Gaps 2;

Qy	2110	GATTCAACATAGACTACACGANTTCATTTCATTTGAAGAGATGATGCCGGAGAAATTTTGTG	2169
Db	1080	GATTCAACATAGACTACCCGANTTCATTTTATTAAGAGATGATGCCGTGAGTATTTTGTG	1021
Qy	2170	TGAAGGGCTTGAACCTCTTTTCACTGTTCCTTATTCAGAGATATTTTGGAAATATATGACT	2229
Db	1020	TGAAGGACTTGAACTGTTTTCATTTGTCTTATTCAGAGATATTTTTCGGAGGTATACGACT	961
Qy	2230	GGAACTCTTAAAGTCTCTTTTGTGAAGACAGCCCTCCCTCGTG-CCCAAGATTTCAATTC	2288
Db	960	GGAACTCTTAAAGTCTCTTTGTTGAAGACAGCCCTCCCTGCTGTCCTCAAGATTTCAATTC	901
Qy	2289	ATGCCACGTTTTGPAAGATTTCTTCCAGATGGAGGAAAGAGTGTGTCCATGCACCAG	2348
Db	900	ATGCCACGTTTTGTAAAGATTTCTTCCAGATGGCGGCAAGAGTGTATCAAATGTACCAG	841
Qy	2349	ATTCTCTCTGT-ACTTGTTAAGTTCAGCAAAAGCCCTGCTGCTGAGAGAGAGATTGCCAA	2407
Db	840	ATTCTCTCTTAAACCCCTGTCCTGTATGACAGGCTTTTGGTCCCTGAGAGAGAGATTGCCAA	781
Qy	2408	TATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAGAGAAATGCAAGATGCAAAAGCATGAT	2467
Db	780	TATGCTCCAAATGGAGAGAGCTCGAGTGGCAGAAATATGCGAAGAGTGCAGAAAGCATGAT	721
Qy	2468	TGTTACCAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATCGTGAACA	2527
Db	720	AGTTACCAACCCCTGGGACGAAACCAAGCCCTGTCCGCATCGATCAACTGATCGTGAGCA	661
Qy	2528	GTTCAACCCCGATGTGATTAATCTTTTCCGATATGTCCTCATCTTTGGGATAAGCCCTGCACA	2587
Db	660	GTTCAACCCCTGAAATGATCACTTTTCCGATATGCTTTCACTTTTGGGATAAGCCCTGCACA	601
Qy	2588	GTTGAGTTATGCAAGGAGACCCACAGTACCAAAACTGTGCAAGAGTTATGTGAAACTTCG	2647
Db	600	GTTGAGTTATGCAAGGAGACCCACAATACCAAAATATATGAAAAGTTTACGTGAAACTTCG	541
Qy	2648	CCACCTCTTAGCAAAATAGTCCCAAAGTCAAAACAACTGACAAACAGAAAGCTGGCACAGAG	2707
Db	540	CCACCTCTTAGCAAAATAGTCCCAAGTCAAAACAGACTGACAAACAGAAAGCTGGCCAGAG	481
Qy	2708	GGAGGAAGCCCTCCAAAAATAAGCGCAGAAAGAAATACAATGAGACGAGAAAGTAAACGTTGGA	2767
Db	480	GGAGGAAGCCCTCCAGAAGATAAGCGCAGAAAGAACACAATGCGTCGAGAAAGTCAACCGTTGA	421
Qy	2768	GCTAAGTAGCCAGGATTTCTGGAAGAACTGCATCCGTTCTGTGATGTCTGTGAGCATGCAT	2827
Db	420	GCTGAGTAGCCAGGATTTCTGGAAGAACTGCATTTCCGTTCTGTGATGTCTGTGAGCATGCAT	361
Qy	2828	GATGCTACCTGTTCTGACCCCATCATATCCGCTACCAAGTGCCTAATGCTAATGATTTTGACAA	2887
Db	360	GATGCTGCCTGTTCTGACCCCATCATATCCGATACCATCAATGCTTAAATGCACTTGACAA	301
Qy	2888	GTTGATAGGATATATCTTTTCCAAGATCGTTGTCTGTTGACGTGGCCATGACTCATCCAAG	2947
Db	300	GTTGATAGGATATATCTTTTTCAGGATCGTTGTCTGCTCCAGCTGGCCATGACTCATCCGAG	241
Qy	2948	TCATCATTTAAATTTTGGAAATGAATCCTGATCATATGCCAGGAATTCATTAATCTAACTGTGG	3007
Db	240	TCATCATTTAAATTTTGGAAATGAATCCTGATCATATGCCAGGAATCTCTGTCTAACTGTGG	181

Qy	3008	AATTGGCAGCCCAAATACGAGACAGAAAGTTTCATC	CATGCACATGCGGAAGNAAGG	3067
Db	180	AATTGGACGCCCAAATATGAGACAGAAAGTTTCATC	CATGCACATGCGGAAGNAAGG	121
Qy	3068	GATTAAACACCTTGATAAATATCATGTCA	CGCTTGGCCAGATGACCCAACTCCCTCGAG	3127
Db	120	AATTAAACACCTTAATAAATATCATGTCA	CGCTCGGCAAGATGCCAACTCCTTCAAG	61
Qy	3128	GATTAAACCAATGAAACGGTTGGAAATCTTCGGGGTGATGCT	TGTTGTAATTTCTGCACGAG	3187
Db	60	GATTAAATCACAATGAAAGGTTGAGTTCTCTGGAGATGCT	GTCTCTGAGTTCTTGACCAAG	1

RESULT 8	
BQ883603	
LOCUS	BQ883603 linear 904 bp mRNA EST 16-AUG-2002
DEFINITION	AGENCOURT 8727466 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6340858
	5', mRNA sequence.
ACCESSION	BQ883603
VERSION	BQ883603.1 GI:22275611
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 904)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2537 row: p column: 11 High quality sequence stop: 695.

```

FEATURES
source
align quality sequence stop: 695.
Location/Qualifiers
1. 904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6340858"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="N1H_MGC_47"
/note="Organ: brain; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

```

ORIGIN

Query Match	17.9%	Score 851.4	DB 5	Length 904
Best Local Similarity	97.8%	Pred. No. 4.8e+209		
Matches 883	Conservative	0	Mismatches 18	Indels 2
				Gaps 2

Note: this is a NIH_MGC Library.

Qy	3480	CTCAATTATCTCTCCACCCGCTCCCAACTTCAAGAGCCAAATCTGATGACAACATTATT	3539
Db	2	CTGGCTCATTTATCCTCTCCCACTCCCACTCAAGAGCC--AATTACTGATGACAACATTATT	60
Qy	3540	GAAACTTCTCCAGTCTTCAAAAAAATTACTGAGTCTTTGAAGAGCAATTTGGAGTAATTTTT	3599
Db	61	GAAACTTCTCCAGTCTTCAAAAAAATTACTGAGTCTTTGAAGAGCAATTTGGAGTAATTTTT	120

QY	3600	ACTCATGTTTCGACTTCTGGCAAGGGCAATTCACATTGAGAACTGTGGGATTTAAACCATCTG	3659
Db	121	ACTCATGTTTCGACTTCTGGCAAGGGCAATTCACATTGAGAACTGTGGGATTTAAACCATCTG	180
QY	3660	ACCTAGGCGCACAAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAGCC	3719
Db	181	ACCTAGGCGCACAAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAGCC	240
QY	3720	ACAGAGTACTTATTCATTATTCATATTCCTCCAGATCATCATGAAGGACATTAATCTTTGGCGA	3779
Db	241	ACAGAGTACTTATTCATTATTCCTCCAGATCATCATGAAGGACATTAATCTTTGGCGA	300
QY	3780	AGCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGAGAGCTGGCATGCAAGAG	3839
Db	301	AGCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGAGAGCTGGCATGCAAGAG	360
QY	3840	TACGCCATACCAACGACAGCAAGAGCGGCTGTGCGCTTCGACCAAGACCTTTGGCG	3899
Db	361	TACGCCATACCAACGACAGCAAGAGCGGCTGTGCGCTTCGACCAAGACCTTTGGCG	420
QY	3900	GACCTTTTGGAAATCATTTATTTAGCAGCGCTGTACACTGATAAGGATTTGGAAATATGTTCA	3959
Db	421	GACCTTTTGGAAATCATTTATTTAGCAGCGCTGTACACTGATAAGGATTTGGAAATATGTTCA	480
QY	3960	ACTTTCATGAATGCTGCTCTTTTCCACGATTTGAAGAAATTCATTTTGAATCAGGATGG	4019
Db	481	ACTTTCATGAATGCTGCTCTTTTCCACGATTTGAAGAAATTCATTTTGAATCAGGATGG	540
QY	4020	ANTGACCCCAANTCCAGCTTCAGCAGTGTGCTTGACACTTAGGACAGAGGAAAGAG	4079
Db	541	AATGACCCCAANTCCAGCTTCAGCAGTGTGCTTGACACTTAGGACAGAGGAAAGAG	600
QY	4080	CCAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCGCCGAACCTAC	4139
Db	601	CCAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCGCCGAACCTAC	660
QY	4140	ACTGTGGCTGTTTATTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGCAACGAATTTTCAG	4199
Db	661	ACTGTGGCTGTTTATTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGCAACGAATTTTCAG	720
QY	4200	CAAGCGAAATGGAGCAGCAATGGATCGCTTGCAAAAATATAATTTTCCCGAGATGCC	4259
Db	721	CAAGCGAAATGGAGCAGCAATGGATCGCTTGCAAAAATATAATTTTCCCGAGATGCC	780
QY	4260	CATCAGAAGCGTTTCATCGCAACGGAAGTACAGACAGAGTTTAAAGAAA-TGAGGTGGGA	4318
Db	781	CATCAGAAGCGTTTCATCGCAACGGAAGTACAGACAGAGTTTANNAGAAATTTGAGGTGGGA	840
QY	4319	AAGAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGAGGGC	4378
Db	841	AAGAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGAGGGC	900
QY	4379	ATG 4381	
Db	901	ATG 903	

RESULT 9
AU120566
LOCUS
DEFINITION AU120566 HEMBB1 Homo sapiens cDNA clone HEMBB1000947 5', mRNA
sequence.
ACCESSION AU120566
VERSION AU120566.2 GI:55778916
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,

TITLE
JOURNAL
COMMENT

Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
On Oct 19, 2000 this sequence version replaced gi:10935801.
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
source

1..896
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBB1000947"
/cissue_type="whole embryo, mainly body"
/dev_stage="embryo, 10 weeks"
/clone_lib="HEMBB1"
/note="Vector: pME18SFL3"

ORIGIN

Query Match	17.2%;	Score 818.8;	DB 1;	Length 896;
Best Local Similarity	97.9%;	Pred. No. 1.4e-200;		
Matches 882;	Conservative 0;	Mismatches 13;	Indels 6;	Gaps 5;
QY	2052	GGATTTTCTATGTTTGCACATGCCCCCTGACCAATATTCACCTGTGTAAAGTAATTAGA	2111	
Db	1	GGATTTTCTATGTTTGCACATGCCCCCTGACCAATATTCACCTGTGTAAAGTAATTAGA	60	
QY	2112	TTCAACATAGACTACACGATTCATTTCAITTAAGAGAGATGATGCCGAGAAATTTTGTGTG	2171	
Db	61	TTCAACATAGACTACACGATTCATTTCAITTAAGAGAGATGATGCCGAGAAATTTTGTGTG	120	
QY	2172	AAAGGGCTTCAAACCTCTTTTTCACCTGTTCCTATTCAGAGATATTTTGGAAATATATGACTGG	2231	
Db	121	AAAGGGCTTGAACCTCTTTTTCACCTGTTCCTATTCAGAGATATTTTGGAAATATATGACTGG	180	
QY	2232	AATCTTTAAAGGCTCTTTTGTGTAAGACAGAGCCCTCCCTGCTGCCAAGATTTTCATTTTCATG	2291	
Db	181	AATCTTTAAAGGCTCTTTTGTGTAAGACAGAGCCCTCCCTGCTGCCAAGATTTTCATTTTCATG	240	
QY	2292	CCAGTTTTTCTAAGATTTCTTTCAGATGGAGGAAAGGAAGTGTGTCCATGCCACAGATT	2351	
Db	241	CCAGTTTTTCTAAGATTTCTTTCAGATGGAGGAAAGGAAGTGTGTCCATGCCACAGATT	300	
QY	2352	CTCTGTACTTGTTTAAGGTGCAGCAAGCCCTGTGTGCTGAGGAGGAGATGCCAATATG	2411	
Db	301	CTCTGTACTTGTTTAAGGTGCAGCAAGCCCTGTGTGCTGAGGAGGAGATGCCAATATG	360	
QY	2412	CTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAGAAGATGCAAGGCGATGTTT	2471	
Db	361	CTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAGAAGATGCAAGGCGATGTTT	420	
QY	2472	ACCAACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2531	
Db	421	ACCAACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	480	
QY	2532	AACCCCGATGTGATTTACTTTTCCGATTTATCGTCCACTTTTGGGATACGCCCTTGCACAGTTG	2591	
Db	481	AACCCCGATGTGATTTACTTTTCCGATTTATCGTCCACTTTTGGGATACGCCCTTGCACAGTTG	540	
QY	2592	AGTTATGAGGAGACCCACAGTACCAAACTGTGGAAAGGATTTATGTGAACCTTCGCCAC	2651	
Db	541	AGTTATGAGGAGACCCACAGTACCAAACTGTGGAAAGGATTTATGTGAACCTTCGCCAC	600	
QY	2652	CTCCTAGCAATATGTGCCAAAGTCAAACTGACAAACAGAACTGGCACAGAGGAG	2711	

```

Db      601  CTCTAGCAATAGTCCCAAGTCAACAACTGACAAACGAGAGCTGGACAGAGGAG 660
Qy      2712 GAAGCCCTCCAAAATACGCGAGAGAAATACATGAGACGAGAGTAACGGTGGAGCTA 2771
Db      661  GAAGCCCTCCAAAATACGCGAGAGAAAT-CAATGAGACGAGAGTAACGGTGGAGCTA 719
Qy      2772 AGTAGCAAGAGTCTCGAAAACCTGGCATCCGTTCTGATGCTCTGAGCATGCAATGATG 2831
Db      720  AGTAGCAAGAGTCTCGAAAACCTGGCATCCGTTCTGATG-CTGTCAATGCAATGATG 778
Qy      2832 CTACTCTTCTGACCCATCATATCCGCTACCAACATGCCTTAATGCATTT-GGACAAAGTT 2890
Db      779  CT--CCTGTCTGACCCATCATATCG--TTCACCAATGCCTATGCAATTTGGGACAAAGTT 835
Qy      2891 GATAGGATATACCTTCCAAAGATCGTTCTGTGTCAGCTGGCCATGATCATCCAAAGTCA 2950
Db      836  GATAGGATATACCTTTCAGAAACCGTTGCTGGTGAACCTGGCCATGATCATCCAAAGTCA 895
Qy      2951 T 2951
Db      896 T 896

```

```

RESULT 10
LOCUS   CR977645
DEFINITION CR977645 RZPD no.9016 Homo sapiens cDNA clone RZPDp901610941 5',
mRNA sequence.
ACCESSION CR977645
VERSION   CR977645.1 GI:68215973
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 826)
Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
and Korn,B.
Human T-Lymphocytes library
Unpublished (2005)
Contact: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp901610941.
RZPDIIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de

```

This clone is available from RZPD;
<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDp901610941>
 contact RZPD (product-support@rzpd.de) for further information.
 Primer name: q3.4 . Primer sequence: CGGATACAAATTCACAG.

FEATURES

```

source
1..826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp901610941"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/notes="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from

```

human T-Lymphocytes with a NotI - oligo(dT) primer [5',
 GACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to SalI adaptors,
 digested with NotI and cloned into the NotI and SalI sites
 of the pQE80LSN_cloned vector"

ORIGIN

```

Query Match      16.7%; Score 797.6; DB 7; Length 826;
Best Local Similarity 99.4%; Pred. No. 4.3e-195;
Matches 811; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy      1305 CGCTCCCAAGTAGGAGAGAGAGAGAGCTCGTTGGGAGCAAGAAAAGACCGTTGGAGT 1364
Db      1 CGTCTCCCAAGTAGGAGAGAGAGAGCTCGTTGGGAGCAAGAAAAGACCGTTGGAGT 60
Qy      1365 GACAACCAAGATTCTTGCAAAAGCAAGAACTATACCTCAATCAAGGAAAAGACCCCGAG 1424
Db      61 GACAACCAAGATTCTTGCAAAAGCAAGAACTATACCTCAATCAAGGAAAAGACCCCGAG 120
Qy      1425 GAGACCATGCTGCACAAGATGAGGAGAGAGAGAACTTCTTAAGCCCTGTGTGATT 1484
Db      121 GAGACCATGCTGCACAAGATGAGGAGAGAGAACTTCTTAAGCCCTGTGTGATT 180
Qy      1485 CGATGCACTCATTTGAGAAAACCTACTCTCCAGTGACCCCATGATGAGTGGGAGATTCT 1544
Db      181 CGATGCACTCATTTGAGAAAACCTACTCTCCAGTGACCCCATGATGAGTGGGAGATTCT 240
Qy      1545 ACAGTGGTTGGAAACGAGTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGG 1604
Db      241 ACAGTGGTTGGAAACGAGTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGG 300
Qy      1605 AGCAGGCAAGAAAAGGCCAAAGCTGCTGGCCCTCCGTTGGGAACTCCAAAGAGAGCTC 1664
Db      301 AGCAGGCAAGAAAAGGCCAAAGCTGCTGGCCCTCCGTTGGGAACTCCAAAGAGAGCTC 360
Qy      1665 GATGAAGATTTAGAGAGTTCACAGTGAATCCGAGTGTGAGTCTCATGAGGAGACAGCCTGT 1724
Db      361 GATGAAGATTTAGAGAGTTCACAGTGAATCCGAGTGTGAGTCTCATGAGGAGACAGCCTGT 420
Qy      1725 TCTAGCAGCTCAGACTCTGAAGTTTTTGTAGCTTATTGAGAAAATCAAAACGAAAAGGCC 1784
Db      421 TCTAGCAGCTCAGACTCTGAAGTTTTTGTAGCTTATTGAGAAAATCAAAACGAAAAGGCC 480
Qy      1785 CACCTGACGAGCTTCATGATGAATTTGTGTACAGATCCAGGCGAGATGAATGATGGA 1844
Db      481 CACCTGACGAGCTTCATGATGAATTTGTGTACAGATCCAGGCGAGATGAATGATGGA 540
Qy      1845 CCACTCTGCAAAATGACGCGCAAGGCAAGCGACAGCAAGGAAATAGGCACAGCATTTATCCT 1904
Db      541 CCACTCTGCAAAATGACGCGCAAGGCAAGCGACAGCAAGGAAATAGGCACAGCATTTATCCT 600
Qy      1905 GGAGAAAGAGGCCATCAAGCCCTGTGCTCTATGACCAAAATCTGTCGAGACTTTTCCAC 1964
Db      601 GGAGAAAGAGGCCATCAAGCCCTGTGCTCTATGACCAAAATCTGTCGAGACTTTTCCAC 660
Qy      1965 TACCGGATCACAGTCTCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAA 2024
Db      661 TACCGGATCACAGTCTCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAA 720
Qy      2025 TACGATGATCACAGATATATCTTTGAAGATTTTCTATGTTGACATG-CCCCCTTGAC 2083
Db      721 TACGATGATCACAGATATATCTTTGAAGATTTTCTATGTTGACATGCCCCCTTGAC 780
Qy      2084 CAATATTCCCTGTGTAAAGTAATTAAGATTCAACAT 2119
Db      781 CAATATTCCCTGTGTAAAGTAATTAAGATTCCACCT 816

```

RESULT 11

```

LOCUS   AU142521
DEFINITION AU142521 Y79A11 Homo sapiens cDNA clone Y79A11000458 5', mRNA
sequence.

```

ACCESSION	AU142521	1	GI:11004042	
VERSION	AU142521.1			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.			
AUTHORS	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.			
TITLE	HRI human cDNA project			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project: 5' - & 3' - end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.			
FEATURES	Location/Qualifiers			
source	1..867			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="Y79AA1000458"			
	/cell_type="retinoblastoma"			
	/cell_line="Y79"			
	/clone_lib="Y79AA1"			
	/note="vector: pME18SFL3"			
ORIGIN				
Query Match	16.7%;	Score	793.8;	DB 1;
Best Local Similarity	96.68;	Pred. No.	4.3e-194;	
Matches	842;	Conservative	0;	Mismatches 25; Indels 5; Gaps 3;
QY	3039	GTTCATCATGCAATGCGGAAGAAAGGATTAACACCTTGATATAATATCATGTACGC	3098	
Db	1	GTTCATCATGCAATGCGGAAGAAAGGATTAACACCTTGATATAATATCATGTACGC	60	
QY	3099	CTTGCCAGATGACCACTCCCTCGAGGATTAACACATGAACGTTGGATTCTGTG	3158	
Db	61	CTTGCCAGATGACCACTCCCTCGAGGATTAACACATGAACGTTGGATTCTGTG	120	
QY	3159	GGTATGCTGTGTGTAATTTCTGACCAAGCTCCATTTGTAATTTGTTCTAGTCTG	3218	
Db	121	GGTATGCTGTGTGTAATTTCTGACCAAGCTCCATTTGTAATTTGTTCTAGTCTG	180	
QY	3219	GAAGAAGAGGATTAGCAACCTATCGGACTGCCATTTGTTCAAGATCAGCCTTGGCATG	3278	
Db	181	GAAGAAGAGGATTAGCAACCTATCGGACTGCCATTTGTTCAAGATCAGCCTTGGCATG	240	
QY	3279	CTAGCAAGAAGAACTTGAACTGATGATTTATGCTGTATGCTCAGCGGCTGACCTTTGT	3338	
Db	241	CTAGCAAGAAGAACTTGAACTGATGATTTATGCTGTATGCTCAGCGGCTGACCTTTGT	300	
QY	3339	AGAGATCGGACCTTCGACATGCAATGCGCAATTTGTTGAAGCTTTAATAGGAGCTGT	3398	
Db	301	AGAGATCGGACCTTCGACATGCAATGCGCAATTTGTTGAAGCTTTAATAGGAGCTGT	360	
QY	3399	TACTTGGAGGGAAGCTCGAGGAAGCCCAAGCAGTTATTGGACGTTTCTCTTTAATGAT	3458	
Db	361	TACTTGGAGGGAAGCTCGAGGAAGCCCAAGCAGTTATTGGACGTTTCTCTTTAATGAT	420	
QY	3459	CCGGAACCTCGGGAAGTCTGGCTCAATTTCTCTCCACCACTCCAACACAGAGCA	3518	
Db	421	CCGGAACCTCGGGAAGTCTGGCTCAATTTCTCTCCACCACTCCAACACAGAGCA	480	
QY	3519	AATACATGATCGACAACTTTATTGAAATTTCTCAGTTCTACAAAACTTACTGAGTTGAA	3578	
Db	481	AATACATGATCGACAACTTTATTGAAATTTCTCAGTTCTACAAAACTTACTGAGTTGAA	540	
QY	3579	GAAGCAATGGAGTAATTTTACTCATGTTTGACCTTCTGGCAAGGGCATTCACATTGAG	3638	
Db	541	GAAGCAATGGAGTAATTTTACTCATGTTTGACCTTCTGGCAAGGGCATTCACATTGAG	600	
QY	3639	ACTGTGGGATTTAACCATCTGACCTTAGCCCAATCAGAGAAATGGAATTCCTAGGTGAC	3698	
Db	601	ACTGTGGGATTTAACCATCTGACCTTAGCCCAATCAGAGAAATGGAATTCCTAGGTGAC	659	
QY	3699	TCATAATGCAACTGCTAGCCACAGAGTACTTTATTCAATTTCCAGATCATCATGAA	3758	
Db	660	TCATAATGCAACTGCTAGCCACAGAGTACTTTATTCAATTTCCAGATCATCATGAA	719	
QY	3759	GGACACTTAACCTTTGGTGGAAAGCTTTTGGTGAATATAGAACTCAGCCCAAGGTAGCG	3818	
Db	720	GGACACTTAACCTTTGGTGGAAAGCTTTTGGTGAATATAGAACTCAGCCCAAGGTAGCG	778	
QY	3819	GAGGAGCTGGCATGCGAGGATGACCAATAACCAAGCAAGACCAAGAGGCTGTGGCG	3878	
Db	779	GAGGAGCTGGCATGCGAGGATGACCAATAACCAAGCAAGACCAAGAGGCTGTGGCG	835	
QY	3879	CTTCGACCAAGACCTTTGGCGGACCTTTTGGGA	3910	
Db	836	GCTTCGACCAAGACCTTTGGCGGACCTTTTGGGA	867	
RESULT 12				
AY411002	1436 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	Mus musculus HCM4078 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	Genomic survey sequence.			
ACCESSION	AY411002			
VERSION	AY411002.1	GI:39766970		
KEYWORDS	GSS.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1 (bases 1 to 1436)			
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios			
JOURNAL	Science 302 (5652), 1960-1963 (2003)			
PUBMED	14671302			
REFERENCE	2 (bases 1 to 1436)			
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA			
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.			
FEATURES	Location/Qualifiers			
source	1..1436			
	/organism="Mus musculus"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:10090"			
	<1..>1436			
	/locus_tag="HCM4078"			
gene				
ORIGIN				
Query Match	16.6%;	Score	793.2;	DB 10;
Best Local Similarity	60.0%;	Pred. No.	7e-194;	
Matches	866;	Conservative	0;	Mismatches 570; Indels 7; Gaps 1;


```
Qy 2928 CTGCCATGACTGATCCAGTCATCATTTAAATTTTGGATGAATCTCTGATCATGCCAGG 2987
Db 1 CTTGCCATGACTCATCCGAGTCACCATTTAAATTTTGGATGAATCTCTGATCATGCCAGG 60
Qy 2988 AATTCAATTTACTCTGGAATTCGGCAGCCCAATACGGAGACAGAAAAGTTTCATCAC 3047
Db 61 AATTCTTTGTTACTCTGGAATTCGGCAGCCCAATATGGAGACAGAAAAGTTTCATCAC 120
Qy 3048 ATGCATGCGGAAAGAAAGGATTAACACCTTGTATAAATATCATGTCAACGCTTGGCCAA 3107
Db 121 ATGCATGCGGAAAGAAAGGATTAACACCTTGTATAAATATCATGTCAACGCTTGGCCAA 180
Qy 3108 GATGACCAACTCCTCGAGATTAACCAATCAACGCTTGGATTCCTGGGTGATGCT 3167
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
Qy 3168 GTTGTGAAATTTCTGACCAAGCGTCAATTTGCTACTATTGTTCTCTAGTCTGGAAGGA 3227
Db 241 GTTGTGAGTTTCTGACCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Qy 3228 GGATTAGCAACTATCGGACTGCAATTTGTCAGAAATCAGCACCTTGCCATGTAGCAAG 3287
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
Qy 3288 AAACCTGAACTGGATCAATTTATGCTGTATGCTCAGCGGCTGACCTTGTAGAGAATCG 3347
Db 361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420
Qy 3348 GACCTTCGACATGCAATGGCCAATTTGTTGAAGCGTTAATAGGAGCTGTTACTTGGAG 3407
Db 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 3408 GGAAGCCTGGAGGAAGCAAGCAGTTATTTTGACGCTTGCTCTTTAATGATCGGACCTG 3467
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 3468 CGGAAGTCTGGCTCAATTAATCTCTCCACCACTCAACCTCAAGAGCCAAATACTGAT 3527
Db 541 CGGAAGTCTGGCTCAATTAATCTCTCCACCACTCAACCTCAAGAGCCAAATAGCGAT 600
Qy 3528 CGCAACTTATGAACTCTCCAGTTCTCAAAATCTTACTGATTTGAAGAGCAAT 3587
Db 601 CGCAACTTATGAACTCTCCAGTTCTCAAGAACTTACGAGTTTGAAGAGCGAT 660
Qy 3588 GGAGTAAATTTTACTCATGTTGCACTTCTGGCAAGGCAATTCACATTTGAGAACTGTGGGA 3647
Db 661 GGAGTATCTTCACTCAGCTCGGCTTCTGGCGAGGCGTTCACTGAGAACCGTGGGC 720
Qy 3648 TTTAACCATCTGACCTTAGGCCCAATCAGAGAAATGGAATTCCTAGTGTGATCATAATG 3707
Db 721 TTTAACCACTGACCTTAGGCCCAATCAGAGAAATGGAATTTCTGGCGGACTCCATAATG 780
Qy 3708 CAATGTTAGCCACAGAGTACTTATTCATTTCTCCAGATCATCATGAGGACACTTA 3767
Db 781 CAGCTGTGCGCCACAGAGTACTTGTTCATTTCTCCGACCATCACGAAGGACACTTG 840
Qy 3768 ACTTTGTTGCGAAGCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGAGCTG 3827
Db 841 ACGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 900
Qy 3828 GGCATGCGAGGATGCGCCATAACCAACGACAGCAAGAGGCGCTGTGGCGCTTCGCACC 3887
Db 901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960
Qy 3888 AAGACCTTGGCGACCTTTGGATCATTTATTGCGCGCTGTACACTGATAAGGATTG 3947
Db 961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020
Qy 3948 GAATATGTTTCATCTTTCATGAATGTCTCTTTCCAGATTGGAAGAAATTCATTTTG 4007
Db 1021 GAATATGTTTCATCTTTCATGAAGAGCTGTCTTCTTTTCCCG-----GAGTTCAATCTG 1073
```

```
Qy 4008 AATCAGGATTTGGAATGACCCCAATCCAGTTTCAGCAGTGTCTTGACACTTAGGACA 4067
Db 1074 AATCAGGATTTGGAACGACCCCAAGTCGCGCTGCAGCAGTGTCTTGACCCCTGAGACA 1133
Qy 4068 GAAGGAAAAGAGCGCAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCGCCATCCAT 4127
Db 1134 GAAGGAAAAGAGCGCTGACATCCCTTATACAAGACTCTGCAGACAGTGGGCGCCATCCAT 1193
Qy 4128 GCCGGAACCTTACACTGTGGCTGTTTATTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGA 4187
Db 1194 GCTAGAACCTTACACTGTGGCTGTTTATTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGA 1253
Qy 4188 CCAAGTATTCAGCAACGGGAATGGGAGCAGCAATGATCGCTTGAATAATATAATTTT 4247
Db 1254 CCAAGCAATTCAGAGCGGAGATGGGAGCAGCAATGATGCACTGGGAAATATAAATTTT 1313
Qy 4248 CCCAGATGCCCATCAGAAAGCGTTTCATCAACGGAAGTACAGACAGAGTAAAAAGAA 4307
Db 1314 CCCAGATGCCCATCAGAAAGCGTTTCATTTGAGCGGAAATACAGACAGAGTAAAGGAA 1373
Qy 4308 ATGAGTGGGAAAAGAGAGCATCAAGAGAGAGAGCGCAGATGAGACTGAAGACATCAAGAA 4367
Db 1374 ATGAGTGGGAAAAGAGAGCATCAGAGAGAGAGCGCGAGGAGCTGAAGACATCAAGAA 1433
Qy 4368 TAA 4370
Db 1434 TAA 1436

RESULT 13
LOCUS CA454899
DEFINITION CA454899 826 bp mRNA linear EST 12-NOV-2002
ACCESSION AGENCOURT_10735747 MAPcL Homo sapiens cdna clone IMAGE:6722379 5',
VERSION CA454899 mRNA sequence.
KEYWORDS CA454899.1 GI:24905082
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 826)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgabbs@mail.nih.gov
Tissue Procurement: Kristi A. Egland, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14285 row: a column: 03
High quality sequence stop: 553.
FEATURES
Location/Qualifiers
1..826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722379"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dt. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Egland, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan. Discovery of new breast
```

ORIGIN	cancer genes encoding membrane and secreted proteins.									
	Manuscript submitted."									
Query Match	16.0%; Score 763.2; DB 6; Length 826;									
	Best Local Similarity 96.6%; Pred. No. 3.7e-186;									
Matches	791; Conservative 0; Mismatches 24; Indels 4; Gaps 1;									
QY	3720	ACAGAGTACTTATTTCATTTCAATTTCCAGATCATCATGAAGACATTAACCTTTTGGCGA	3779							
DB	1	ACAGAGTACTTATTTCATTTCCAGATCATCATGAAGACATTAACCTTTTGGCGA	60							
QY	3780	AGCTCTTTGGTGAATTAATGAATCTCAGCCCAAGGTAGCGAGAGCTGGCATGCGAGG	3839							
DB	61	AGCTCTTTGGTGAATTAATGAATCTCAGCCCAAGGTAGCGAGAGCTGGCATGCGAGG	120							
QY	3840	TACGCCATACCAACGACAGACCAAGAGCCCTGTGCGCTTCGACCAAGACCTTTGGCG	3899							
DB	121	TACGCCATACCAACGACAGACCAAGAGCCCTGTGCGCTTCGACCAAGACCTTTGGCG	180							
QY	3900	GACCTTTTGGAAATCAATTTATTGAGCGCTGTACACTGATGAAGATTTGGAAATGTTTCAT	3959							
DB	181	GACCTTTTGGAAATCAATTTATTGAGCGCTGTACACTGATGAAGATTTGGAAATGTTTCAT	240							
QY	3960	ACTTTCATGAATGTCTGCTTTTCCACGATTTGAAGAAATTCATTTTGAATCAGGATGG	4019							
DB	241	ACTTTCATGAATGTCTGCTTTTCCACGATTTGAAGAAATTCATTTTGAATCAGGATGG	300							
QY	4020	ANTGACCCCAATCCAGCTTCAGCAGTGTGCTTGACACTTAGGACAGAAGAAAGAG	4079							
DB	301	ANTGACCCCAATCCAGCTTCAGCAGTGTGCTTGACACTTAGGACAGAAGAAAGAG	360							
QY	4080	CCAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCCATCCCATGCCGCAACCTAC	4139							
DB	361	CCAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCCATCCCATGCCGCAACCTAC	420							
QY	4140	ACTGTGGCTGTTTATTTCAAGGGGAGAAAGATAGGCTGTGGGAAAGACCAAGTATTCAG	4199							
DB	421	ACTGTGGCTGTTTATTTCAAGGGGAGAAAGATAGGCTGTGGGAAAGACCAAGTATTCAG	480							
QY	4200	CAAGCGAAATGGGACAGCAATGGATCGCTTGAAATAATATTTTCCCGAGATGCC	4259							
DB	481	CAAGCGAAATGGGACAGCAATGGATCGCTTGAAATAATATTTTCCCGAGATGCC	540							
QY	4260	CATCAGAAGCGGTTTCATCGAACGNAAGTACAGACAGAGTTTAAAGAAATGAGTGGGAA	4319							
DB	541	CATCAGAAGCGGTTTCATCGAACGNAAGTACAGACAGAGTTTAAAGAAATGAGTGGGAA	600							
QY	4320	AGAGAGCATCAAGAGAGAGAGCCAGATGAGCTGAAGACATCAAGAAATTAAGAGGGCA	4379							
DB	601	AGAGAGCATCAAGAGAGAGAGCCAGATGAGCTGAAGACATCAAGAAATTAAGAGGGCA	660							
QY	4380	TGCAAGTGTGAGATTTACTTGTCTCAGTACTGTGCTATTTGAGACCTAGCC	4439							
DB	661	TGCAAGTGTGAGATTTACTTGTCTCAGTACTGTGCTATTTGAGACCTAGCC	720							
QY	4440	TAGTTTTCTCAGACAAATGAACGAAGTGTCTCATTTGAAATAAATACAGAGTCAAAATC	4499							
DB	721	TAGTTTTCTCAGACAAATGAACGAAGTGTCTCATTTGAAATAAATACAGAGTCAAAATC	780							
QY	4500	GC-----TATTGTGTTTAAATGATCTGTTTATTAGCTGGA	4534							
DB	781	GCCTATTGGTGGGTTAAAGGAACCGGTTTTTATAGCTGGA	819							
RESULT 14										
BU155281										
LOCUS										
DEFINITION										
AGENCOURT 7944703 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6148714										
5', mRNA sequence.										
ACCESSION										
BU155281										
VERSION										
BU155281.1 GI:22668813										
KEYWORDS										
EST.										

SOURCE	Homo sapiens (human)									
ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini, Hominidae; Homo.									
REFERENCE	1 (bases 1 to 888)									
AUTHORS	NIH-MGC http://mgs.nci.nih.gov/ .									
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)									
JOURNAL	Unpublished (1999)									
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue procurement: ATCC/DCTD/DTP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13480 row: j column: 11 High quality sequence stop: 669.									
FEATURES	Location/Qualifiers									
source	1..888 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6148714" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 72" /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."									
ORIGIN	Query Match 15.9%; Score 755.8; DB 5; Length 888; Best Local Similarity 96.8%; Pred. No. 3.1e-184; Matches 814; Conservative 0; Mismatches 22; Indels 5; Gaps 4;									
QY	1	CTGCTCTTGTGTACCTCGGGTAGTAGCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG	60							
DB	2	CTGCTCTTGTGTACCTCGGGTAGTAGCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG	61							
QY	61	AGCCTTTTATAGTGTCTTTTCCGGGATGTCAAGATACAGAAATGACTGTGAATCAA	120							
DB	62	AGCCTTTTATAGTGTCTTTTCCGGGATGTGAAGATACAGAAATGACTGTGAATCAA	121							
QY	121	CCCATATCATCAAGGAGCTGATAATCTAGTGGAGAGTTAGACGTGTGCATATCTCACTA	180							
DB	122	CCCATATCATCAAGGAGCTGATAATCTAGTGGAGAGTTAGACGTGTGCATATCTCACTA	181							
QY	181	TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCGGA	240							
DB	182	TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCGGA	241							
QY	241	ACATCATGATGACAGGGAACACATGTCAAGATGTGTTCCACCCGGGACGAGGGGTC	300							
DB	242	ACATCATGATGACAGGGAACACATGTCAAGATGTGTTCCACCCGGGACGAGGGGTC	301							
QY	301	CCCGAGGACGAGGAGGACATGGAGCCAGACCTCAGCACCATCTCTTTAGGCCCCCAAAATC	360							
DB	302	CCCGAGGACGAGGAGGACATGGAGCCAGACCTCAGCACCATCTTTAGGCCCCCAAAATC	361							
QY	361	TGAGGCTGCTTCAACCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC	420							
DB	362	TGAGGCTGCTTCAACCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC	421							
QY	421	CTTCCACACCTTTCTCAAACTCTCAGCCCCCAATTTTCCCTCCACGACGACGACTTTG	480							
DB	422	CTTCCACACCTTTCTCAAACTCTCAGCCCCCAATTTTCTCCCTCCACGACGACGACTTTG	481							
QY	481	TACCTTTCCCCCACCACCATGCTCCGTCCGTCAGCGGCAAGGCCCTCTTCCCCCTTGCCCAATCA	540							

```

Db      482 TACCCTTCCGCCACCCATGCTCGTCAAGGCAAGGCCCTCTTCCGCCCTGCCCCAATCA 541
Qy      541 GCGCGCTTCCCAACACACAGATGAGGACCCCTTCAGTTCCTCTGTTTCTC 600
Db      542 GCGCGCTTCCCAACACACAGATGAGGACCCCTTCAGTTCCTCTGTTTCTC 601
Qy      601 CCATGCCACCAATGCTTGTCTAATAACCCCCAGTCCCTGCGGCACTCTCTGGAC 660
Db      602 CCATGCCACCAATGCTTGTCTAATAACCCCCAGTCCCTGCGGCACTCTCTGGAC 661
Qy      661 AAGGCACTTTCCTTCATGATG-CCGCCCTCCCTCCATGCCCTCATCCCGCGCCCTCCA 719
Db      662 AAGGCACTTTCCTTCATGATG-CCGCCCTCCCTCCATGCCCTCATCCCGCGCCCTCCA 721
Qy      720 GTATGCCGAGCAGGTTAATATCAGTACCTCC-GGGCTATTCTACCAACTT-CC 777
Db      722 GTATGCCGAGCAGGTTAATATCAGTACCTCCGGGGCTATTCTACCAACTTCCC 781
Qy      778 CACCTCCAGTTCCTT- AATAGTTTCCAGAACACCTAGTTCCTTCCTGCGCGCTAA 835
Db      782 CACCTCCAGTTCCTTAAAGTTTCCAGAACACCTAGTTCCTTTCCTGCGCGCAG 841
Qy      836 T 836
Db      842 T 842

RESULT 15
BI753613
LOCUS   603028687F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5199165 5',
DEFINITION mRNA sequence.
ACCESSION BI753613
VERSION    1
KEYWORDS  BI753613.1 GI:15745191
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 794)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-remail.nih.gov
         Tissue Procurement: Life Technologies, Inc.
         cDNA Library Preparation: Life Technologies, Inc.
         cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Incyte Genomics, Inc.
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: L14M11498 row: m column: 22
         High quality sequence stop: 792.
         Location/Qualifiers
             1..794
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:5199165"
                 /lab_host="DH10B"
                 /clone_lib="NIH_MGC_114"
                 /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
                 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
                 male brains, age range 23-27 yo. Library is oligo-dT
                 primed and directionally cloned (EcoRV site is destroyed
                 upon cloning). Average insert size 1.5 kb, insert size
                 range 1-3 kb. Library is normalized and enriched for
                 full-length clones and was constructed by C. Gruber
                 (Invitrogen). Research Genetics tracking code 019. Note:
                 this is a NIH_MGC Library."

FEATURES
source

```

```

Query Match      15.8%; Score 752.6; DB 3; Length 794;
Best Local Similarity 99.1%; Pred. No. 2.le-183;
Matches 788; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

Qy      3309 ATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATCAANTGGCC 3368
Db      1 ATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATCAANTGGCC 59
Qy      3369 AATTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGCTCGAGGAAGCCAAAG 3428
Db      60 AATTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGCTCGAGGAAGCCAAAG 119
Qy      3429 CAGTTATTGAGACGCTTGCTCTTTAATGATCGGACCTCGCGGAAGTCTGGCTCAATTAT 3488
Db      120 CAGTTATTGAGACGCTTGCTCTTTAATGATCGGACCTCGCGGAAGTCTGGCTCAATTAT 179
Qy      3489 CCTCTCCACCCACATCCAACTACAGAGCAAAATCTGATCGACAACTTATTGAAAATCTCT 3548
Db      180 CCTCTCCACCCACATCCAACTACAGAGCAAAATCTGATCGACAACTTATTGAAAATCTCT 239
Qy      3549 CCAGTTCTACAAAACCTTACTGAGTTTGAAGAAGCAATTTGGAGTAATTTTACTCATGTT 3608
Db      240 CCAGTTCTACAAAACCTTACTGAGTTTGAAGAAGCAATTTGGAGTAATTTTACTCATGTT 299
Qy      3609 CGACTTCTGGCAAGGCGATTACATTTGAGAACTGTGGGATTTAACCATCTGACCCCTAGGC 3668
Db      300 CGACTTCTGGCAAGGCGATTACATTTGAGAACTGTGGGATTTAACCATCTGACCCCTAGGC 359
Qy      3669 CACAATCAGAGAAATGGAAATTCCTTAGTGACTCCATAATCAACTGGTAGCCACAGAGTAC 3728
Db      360 CACAATCAGAGAAATGGAAATTCCTTAGTGACTCCATAATCAACTGGTAGCCACAGAGTAC 419
Qy      3729 TTATTTCATTCATTTCCAGATCATCATGAAGGACACTTAACTTTGTCGAAGCTCTTTG 3788
Db      420 TTATTTCATTCATTTCCAGATCATCATGAAGGACACTTAACTTTGTCGAAGCTCTTTG 479
Qy      3789 GTGAATTAATAGAACTCAGGCCAAGGTAGCGGAGAGCTGGGCATGCGAGGATAGCCATA 3848
Db      480 GTGAATTAATAGAACTCAGGCCAAGGTAGCGGAGAGCTGGGCATGCGAGGATAGCCATA 539
Qy      3849 ACCAACGACAGACCAAGAGGCTGTGCGCTTCGCACCAAGACCTTGGCGGACCTTTTG 3908
Db      540 ACCAACGACAGACCAAGAGGCTGTGCGCTTCGCACCAAGACCTTGGCGGACCTTTTG 599
Qy      3909 GAATCATTTATTGACGGCTGTACACTGATAAGGATTTGGAATATGTTTCACTTTTATG 3968
Db      600 GAATCATTTATTGACGGCTGTACACTGATAAGGATTTGGAATATGTTTCACTTTTATG 659
Qy      3969 AATGCTGCTTCTTTTCCAGATTGAAATTCATTTTGAATCAGGATTTGGAATGACCCC 4028
Db      660 AATGCTGCTTCTTTTCCAGATTGAAATTCATTTTGAATCAGGATTTGGAATGACCCC 719
Qy      4029 AATCCCGAGCTTCAGCAGTGTGCTTGACAC-TTAGGACAGAGGAA-AAAGCCAGACA 4086
Db      720 AATCCCGAGCTTCAGCAGTGTGCTTGACAC-TTAGGACAGAGGAA-AAAGCCAGACA 779
Qy      4087 TTCTCTGTACAGA 4101
Db      780 TTCTCTGTACAGA 794

```

Search completed: December 27, 2005, 03:11:38
Job time : 17368 secs

ORIGIN

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 15:06:28 ; Search time 775 Seconds

(without alignments)
10926.844 Million cell updates/sec

Title: US-10-774-974-1
Perfect score: 4764
Sequence: 1 cgtcttgtagctcggtgta.....aaaaaaaaaaaaaaaaaaaa 4764

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1/1.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4764	100.0	4764	3	US-09-900-425B-1
2	370	7.8	397	3	US-09-513-999C-558
3	259.6	5.4	354	3	US-09-513-999C-8294
4	194.6	4.1	430	3	US-09-513-999C-11432
5	92	1.9	7218	2	US-08-232-463-14
6	65	1.4	7218	2	US-08-232-463-14
7	62.8	1.3	726	3	US-09-107-532A-1780
8	62.8	1.3	732	3	US-08-965-087-1
9	62.4	1.3	6251	3	US-08-956-171E-25
10	62.4	1.3	6251	3	US-08-781-986A-25
11	61.6	1.3	320	3	US-09-165-264-13
12	61	1.3	699	3	US-09-583-110-1893
13	61	1.3	711	3	US-09-107-433-1791
14	60	1.3	687	3	US-09-543-681A-3474
15	59.8	1.3	320	3	US-09-165-264-11
16	59.4	1.2	699	2	US-08-869-674-1
17	59.4	1.2	699	3	US-09-213-010-1
18	59.4	1.2	699	3	US-09-213-011-1
19	59.4	1.2	1830121	3	US-09-557-884-1
20	59.4	1.2	1830121	3	US-09-643-990A-1
21	59.4	1.2	1830121	3	US-10-158-865-1
22	59	1.2	319	3	US-09-165-264-8
23	58	1.2	318	3	US-09-165-264-12
24	57.8	1.2	13121	3	US-08-961-527-126

C	25	57.6	1.2	320	3	US-09-165-264-14	Sequence 14, Appl
C	26	57.6	1.2	762	3	US-09-252-991A-12924	Sequence 12924, A
C	27	57.6	1.2	1011	3	US-09-252-991A-12560	Sequence 12560, A
C	28	57.4	1.2	320	3	US-09-165-264-7	Sequence 7, Appl
		56	1.2	741	3	US-09-134-001C-1950	Sequence 1950, Ap
		54.8	1.2	53526	3	US-08-658-136-2	Sequence 2, Appl
		54.8	1.2	53577	3	US-08-658-136-1	Sequence 1, Appl
		54.8	1.2	53577	3	US-08-460-215A-1	Sequence 1, Appl
		54.4	1.1	738	3	US-09-710-279-1657	Sequence 1657, Ap
C	34	54.4	1.1	3649	3	US-09-438-185A-1	Sequence 1, Appl
C	35	54.4	1.1	1230230	3	US-09-949-016-12160	Sequence 12160, A
		54	1.1	119032	3	US-09-949-016-17268	Sequence 17268, A
C	37	54	1.1	1230025	3	US-09-949-016-16420	Sequence 16420, A
C	39	53.8	1.1	234884	3	US-09-790-988-1	Sequence 1, Appl
C	40	53.8	1.1	640681	3	US-09-553-863-17	Sequence 17, Appl
		52.6	1.1	762	3	US-09-949-016-16775	Sequence 16775, A
		52	1.1	12695	3	US-09-489-039A-4378	Sequence 4378, Ap
		51.2	1.1	684	3	US-09-918-909A-27	Sequence 27, Appl
		51.2	1.1	1738	3	US-09-634-238-203	Sequence 203, App
		51.2	1.1	1753	3		

ALIGNMENTS

RESULT 1
US-09-900-425B-1
; Sequence 1, Application US/09900425B
; Patent No. 6737512
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISI5029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425B
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 4764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-900-425B-1

Query Match	100.0%	Score 4764;	DB 3;	Length 4764;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 4764;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CTGTCTTGTACTGCGGTAGTAGCTTGGCTTGTCTGTACGGCGATCTCGCGCCCGAG	60	
DB	1	CTGTCTTGTACTGCGGTAGTAGCTTGGCTTGTCTGTACGGCGATCTCGCGCCCGAG	60	
QY	61	AGCCTTTTATAGTGTCTTTCCCGGGGATGCAAGGATACAGAAATGACTGTGAATCAA	120	
DB	61	AGCCTTTTATAGTGTCTTTCCCGGGGATGCAAGGATACAGAAATGACTGTGAATCAA	120	
QY	121	CCCATATCATCAAGGAGCTGATAATCTAGTGGAGAGTTAGACGTGTGCATCTTACTA	180	
DB	121	CCCATATCATCAAGGAGCTGATAATCTAGTGGAGAGTTAGACGTGTGCATCTTACTA	180	
QY	181	TGATATGAGCGAGCTCTCTGAGCTTATTTCTCTGTGGAAGATGTGACATATCCAGCGGA	240	
DB	181	TGATATGAGCGAGCTCTCTGAGCTTATTTCTCTGTGGAAGATGTGACATATCCAGCGGA	240	
QY	241	ACATCATGATGACAGGAAACACATGTCAGATGTCCTTCCACCGGACGAGGGGTC	300	
DB	241	ACATCATGATGACAGGAAACACATGTCAGATGTCCTTCCACCGGACGAGGGGTC	300	
QY	301	CCGAGGACGAGGAGCATGAGGACGAGCCCTCAGCACCATCTTTAGGCCCAAAATC	360	
DB	301	CCGAGGACGAGGAGCATGAGGACGAGCCCTCAGCACCATCTTTAGGCCCAAAATC	360	

Db 2521 GTGAACAGTTCAACCCCGATGTGATTTACTTTTCCGATTATCGTCCACTTTGGGATAGCC 2580
Qy 2581 CTGCAAGTTGAGTTATGACGAGAGACCCACAGTACCAAAAACTGTGTGGAAGAGTTATGTGA 2640
Db 2581 CTGCAAGTTGAGTTATGACGAGAGACCCACAGTACCAAAAACTGTGTGGAAGAGTTATGTGA 2640
Qy 2641 AACTTCGCCACCTCTAGCAAAATAGTCCCAAAAGTCAAACTGCAAAACAGAGCTGG 2700
Db 2641 AACTTCGCCACCTCTAGCAAAATAGTCCCAAAAGTCAAACTGCAAAACAGAGCTGG 2700
Qy 2701 CACAGAGGAGGAAGCCCTCCAAAAATACGGCAGAGAATACAATGACAGCAGAAAGTAA 2760
Db 2701 CACAGAGGAGGAAGCCCTCCAAAAATACGGCAGAGAATACAATGACAGCAGAAAGTAA 2760
Qy 2761 CGGTGAGCTAAGTAGCAAGGATCTGGAAAACTGGCAATCTGGCATCTGATGTCTGTGTCAGC 2820
Db 2761 CGGTGAGCTAAGTAGCAAGGATCTGGAAAACTGGCAATCTGGCATCTGATGTCTGTGTCAGC 2820
Qy 2821 ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCCCTAATGCAT 2880
Db 2821 ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCCCTAATGCAT 2880
Qy 2881 TGGACAAGTTGATAGGATATACCTTTCCAAAGATGTTTCTGTCTGTGTCAGCTGGCCATGACTC 2940
Db 2881 TGGACAAGTTGATAGGATATACCTTTCCAAAGATGTTTCTGTCTGTGTCAGCTGGCCATGACTC 2940
Qy 2941 ATCCAAAGTCAATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGACGAAATTCATATCTA 3000
Db 2941 ATCCAAAGTCAATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGACGAAATTCATATCTA 3000
Qy 3001 ACTGTGGAATTCGGACGCCCAATACGGAGACAGAAAGTTTCATCATGCAATGCCGGA 3060
Db 3001 ACTGTGGAATTCGGACGCCCAATACGGAGACAGAAAGTTTCATCATGCAATGCCGGA 3060
Qy 3061 AGAAAGGATTAACACTTTGATATAATCATGTGACGCTTGGCCAAAGATGACCAACTC 3120
Db 3061 AGAAAGGATTAACACTTTGATATAATCATGTGACGCTTGGCCAAAGATGACCAACTC 3120
Qy 3121 CCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTGTGGAATTC 3180
Db 3121 CCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTGTGGAATTC 3180
Qy 3181 TGACGAGCTCCATTTGTACTATTGTTTCTAGTCTGGAAGAGAGGATTAGCAACT 3240
Db 3181 TGACGAGCTCCATTTGTACTATTGTTTCTAGTCTGGAAGAGAGGATTAGCAACT 3240
Qy 3241 ATCGGACTGCCATTTGTTGAGATCAGACCTTGCATGCTAGCAAGAAACTTTGAAGCTG 3300
Db 3241 ATCGGACTGCCATTTGTTGAGATCAGACCTTGCATGCTAGCAAGAAACTTTGAAGCTG 3300
Qy 3301 ATCCATTTATGCTGATGCTCACGGGCTGACCTTTGTAGAGAACTCGGACCTTCGACATG 3360
Db 3301 ATCCATTTATGCTGATGCTCACGGGCTGACCTTTGTAGAGAACTCGGACCTTCGACATG 3360
Qy 3361 CAAATGCCAATGTTTGAAGCTTAATAGGAGCTGTTTACTTTGGAGGAGAGCTCGAGG 3420
Db 3361 CAAATGCCAATGTTTGAAGCTTAATAGGAGCTGTTTACTTTGGAGGAGAGCTCGAGG 3420
Qy 3421 AAGCCAAAGCAGTTATTTGAGAGCTGCTCTTTAATGATCGGACCTCGCGAAGTCTGGC 3480
Db 3421 AAGCCAAAGCAGTTATTTGAGAGCTGCTCTTTAATGATCGGACCTCGCGAAGTCTGGC 3480
Qy 3481 TCAATTTATCCTCTCCACCACTCAAGAGCAAAATCTGATCGCAAACTTTATTG 3540
Db 3481 TCAATTTATCCTCTCCACCACTCAAGAGCAAAATCTGATCGCAAACTTTATTG 3540
Qy 3541 AAACCTCTCCAGTTCTACAAAACTTACTGAGTTTGAAGAGCAATTTGGAGTAAATTTTA 3600
Db 3541 AAACCTCTCCAGTTCTACAAAACTTACTGAGTTTGAAGAGCAATTTGGAGTAAATTTTA 3600
Qy 3601 CTCATGTTGAGCTCTGGCAAGGCAATTCATTTGAGAACTGCGGATTTAACCTCTGA 3660
Db 3601 CTCATGTTGAGCTCTGGCAAGGCAATTCATTTGAGAACTGCGGATTTAACCTCTGA 3660

Qy 3661 CCTAGGCCACAAATCAGAGAAATGGAATTTCTAGTGACTCCATAATGCAACTGGTAGCCA 3720
Db 3661 CCTAGGCCACAAATCAGAGAAATGGAATTTCTAGTGACTCCATAATGCAACTGGTAGCCA 3720
Qy 3721 CAGAGTACTTATTATTCAATTTCCAGATCATCATGAAGGACACTTAACTTTGTGTGCAA 3780
Db 3721 CAGAGTACTTATTATTCAATTTCCAGATCATCATGAAGGACACTTAACTTTGTGTGCAA 3780
Qy 3781 GCTTTTGTGTAATAATAGAACTCAGGCCAAGTGTAGCGGAGGAGCTGGGCAATGAGAGT 3840
Db 3781 GCTTTTGTGTAATAATAGAACTCAGGCCAAGTGTAGCGGAGGAGCTGGGCAATGAGAGT 3840
Qy 3841 AGCCATACCAACGACAGCAAGAGGCTGTGGCGCTTCGCACCAAGACCTTGGCGG 3900
Db 3841 AGCCATACCAACGACAGCAAGAGGCTGTGGCGCTTCGCACCAAGACCTTGGCGG 3900
Qy 3901 ACCTTTTGGAAATCATTTATTGACGCTGTACACTGATGAAGATTTGGAAATATGTTTCA 3960
Db 3901 ACCTTTTGGAAATCATTTATTGACGCTGTACACTGATGAAGATTTGGAAATATGTTTCA 3960
Qy 3961 CTTTCATGAATGTCTGCTTTTCCAGATTTGAAAGATTTCAATTTGAAATCAGGATTTGA 4020
Db 3961 CTTTCATGAATGTCTGCTTTTCCAGATTTGAAAGATTTCAATTTGAAATCAGGATTTGA 4020
Qy 4021 ATGACCCCAATCCAGCTTCAGCAGTGTGCTTGAACCTTTAGACAGAGGAAAGAGC 4080
Db 4021 ATGACCCCAATCCAGCTTCAGCAGTGTGCTTGAACCTTTAGACAGAGGAAAGAGC 4080
Qy 4081 CAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTACA 4140
Db 4081 CAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGAACTTACA 4140
Qy 4141 CTGTGCTGTTTATTTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Db 4141 CTGTGCTGTTTATTTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Qy 4201 AAGCGAAATGGGAGCAGCAATGGATGCGCTTGAATAATATATTTTCCACAGATGGCC 4260
Db 4201 AAGCGAAATGGGAGCAGCAATGGATGCGCTTGAATAATATATATTTTCCACAGATGGCC 4260
Qy 4261 ATCAGAAGCGGTTTCATCGAAACGGAAGTACAGACAGAGTTTAAAGAAATGAGGTGGGAAA 4320
Db 4261 ATCAGAAGCGGTTTCATCGAAACGGAAGTACAGACAGAGTTTAAAGAAATGAGGTGGGAAA 4320
Qy 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGATGAGACTGAAGACATCAAGAAATTAAGGAGGCA 4380
Db 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGATGAGACTGAAGACATCAAGAAATTAAGGAGGCA 4380
Qy 4381 GCAAGTGTGGAGTATTTTACTTGTCTCAGTAACTGTGACTGTTGTCTATTGAGACTAGCCT 4440
Db 4381 GCAAGTGTGGAGTATTTTACTTGTCTCAGTAACTGTGACTGTTGTCTATTGAGACTAGCCT 4440
Qy 4441 AGTTTTCTTCAGACAAATGAACGAAGTGTCTCATTTGAAATAAATAACAGAGTCAAAATCG 4500
Db 4441 AGTTTTCTTCAGACAAATGAACGAAGTGTCTCATTTGAAATAAATAACAGAGTCAAAATCG 4500
Qy 4501 CTATTGTTGTTTAAATGATCTGTTTAACTGATGCTTTTATTTACAAAGTATTAGATT 4560
Db 4501 CTATTGTTGTTTAAATGATCTGTTTAACTGATGCTTTTATTTATCAAAAGTATTAGATT 4560
Qy 4561 TTTCTTTATTTAAACGGAATACTTGACTTTGGTAATGTGCAATTAATTCCTTTTATTTTG 4620
Db 4561 TTTCTTTATTTAAACGGAATACTTGACTTTGGTAATGTGCAATTAATTCCTTTTATTTTG 4620
Qy 4621 CTCCTTTAAATAATAAATTAAGAGCAATTTCTATGTGGAATAGATCTGTGTTTTCCAT 4680
Db 4621 CTCCTTTAAATAATAAATTAAGAGCAATTTCTATGTGGAATAGATCTGTGTTTTCCAT 4680
Qy 4681 CTGTGTCCAGATTTGACCCCTAGACTTTTCAATTTGAAGAGTAAAAAATTTGACTTTACTAG 4740
Db 4681 CTGTGTCCAGATTTGACCCCTAGACTTTTCAATTTGAAGAGTAAAAAATTTGACTTTACTAG 4740

QY 4741 TAAAAAAAAAAAAAAAAAAAAA 4764
Db 4741 TAAAAAAAAAAAAAAAAAAAAA 4764

RESULT 2
US-09-513-999C-558
; Sequence 558, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961

; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 558
; LENGTH: 397

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..397
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 53..121
; OTHER INFORMATION: score 5.9
; OTHER INFORMATION: seq TLKPLTLQKSN/KR
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 325
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 3
; OTHER INFORMATION: Xaa=Ala or Pro

US-09-513-999C-558

Query Match 7.8%; Score 370; DB 3; Length 397;
Best Local Similarity 98.5%; Pred. No. 9.1e-95;
Matches 392; Conservative 2; Mismatches 2; Indels 2; Gaps 2;
QY 1655 GAGCAAGCTCGATCAAGATTAGAGAGTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGA 1714
Db 1 GAGCAAGCTCGATCAAGATTAGAGAGTTCCAGTGAATCCGA-TGTGAGTCTGATGATGA 59
QY 1715 CAGCACCTGTTCTAGCAGCTCAGACTCTGAAGTTTTTGAGGTTATTGCGAAGATCAACG 1774
Db 60 CAGCACCTGTTCTAGCAGCTCAGACTCTGAAGTTTTTGAGGTTATTGCGAAGATCAACG 119

QY 1775 CAAAAGGCC-ACCCTGACCGACTTCATGATGAATTTGCTACAACGATCCAGGCCAGA 1833
Db 120 CAAAAGGCCAACCTTGACCGACTTCATGATGAATTTGCTACAACGATCCAGGCCAGA 179
QY 1834 TGAATGATGACCACTCTGCAAAATGCAAGCGCAAGGCAAGCAGCAGCAAGGAATTAGGCACA 1893
Db 180 TGAATGATGACCACTCTGCAAAATGCAAGCGCAAGGCAAGCAGCAGCAAGGAATTAGGCACA 239
QY 1894 GCATTATCTGGAGAGAGCCCATCAAGCCCTGTGCTCTATGACCAACAATGCTGGCA 1953
Db 240 GCATTATCTGGAGAGAGCCCATCAAGCCCTGTGCTCTATGACCAACAATGCTGGCA 299
QY 1954 GACTTTTCCACTACCGGATCACAGTCTCCCGCCTACGAACCTTTTAACTGACAGGCCAA 2013

Db 300 GACTTTTCCACTACCGGATCACAGTWTCCCGCTACGAACCTTTTAACTGACAGGCCAA 359
QY 2014 CTGTTATAGAAATACGATGATCAGGATATATCTTTGAA 2051
Db 360 CTGTTATAGAAATACGATGATCAGGATATATCTTTGAA 397

RESULT 3
US-09-513-999C-8294
; Sequence 8294, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 8294
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 339
; OTHER INFORMATION: r=a or g
US-09-513-999C-8294

Query Match 5.4%; Score 259.6; DB 3; Length 354;
Best Local Similarity 99.3%; Pred. No. 2.5e-63;
Matches 270; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 CTGCTTGTGTAAGCTGCGGTAGTAGCTGCTGCTGCGCGGATCTGCGCGGCGGAG 60
Db 84 CTGCTTGTGTAAGCTGCGGTAGTAGCTGCTGCGCGGATCTGCGCGGCGGAG 143
QY 61 AGCCTTTTATAGCTTCTTTCCCGGGGATGTGAAGATACAGAAATGACTGTGAATCAA 120
Db 144 AGCCTTTTATAGCTTCTTTCCCGGGGATGTGAAGATACAGAAATGACTGTGAATCAA 203
QY 121 CCCATATCATCAAGGAGCTGATTAATCTAGTGGAGAGTTAGACGTGTGCATATTCACTA 180
Db 204 CCCATATCATCAAGGAGCTGATTAATCTAGTGGAGAGTTAGACGTGTGCATATTCACTA 263
QY 181 TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCCGA 240
Db 264 TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCCGA 323
QY 241 ACATCATGATGCAAGGAAACACATGTCAACA 272
Db 324 ACATCATGATGCA-GGAAACATGTCAACA 354

RESULT 4
US-09-513-999C-11432
; Sequence 11432, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24

```
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 11432
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 415
; OTHER INFORMATION: r=a or g
US-09-513-999C-11432

Query Match      4.1%; Score 194.6; DB 3; Length 430;
Best Local Similarity 96.7%; Pred. No. 9.9e-45;
Matches 208; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 58 GAGAGCCCTTTATAGTTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAAT 117
DB 217 GAGTTACTTTTCAGGTTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAAT 276
QY 118 CAACCCATATCATCAAGGAGCTGATATCTAGTGGAAAGATTAGACGTGTGCATCTTCA 177
DB 277 CAACCCATATCATCAAGGAGCTGATATCTAGTGGAAAGATTAGACGTGTGCATCTTCA 336
QY 178 CTATGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGACATATCCAGGC 237
DB 337 CTATGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGACATATCCAGGC 396
QY 238 GGAACATCATGATGCGAGGAAACACATGTCCACAGA 272
DB 397 GGAACATCATGATGCGA-GGAACACATGTCCACAGA 430

RESULT 5
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109

; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match      1.9%; Score 92; DB 2; Length 7218;
Best Local Similarity 2.1%; Pred. No. 1.2e-14;
Matches 8; Conservative 258; Mismatches 118; Indels 0; Gaps 0;

QY 419 CCCTTCCACCACTTCTCAAACTCTCCAGCCCAATTTTCTCCTCCACGACGACTT 478
DB 1067 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1126
QY 479 TGTACCCCTTCCCCCAACCCATGCTCGTCAAGGCAAGCCCTTCCGCCCTGCCCAAT 538
DB 1127 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1186
QY 539 CAGGCGCCTTTCACCAACACAGATGAGGACCCCTTCCAGTTCTCTGTTTCC 598
DB 1187 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1246
QY 599 TCCATGCCACCAACCAATGCTTGTCTTAATAACCCCAAGTCCCTGGGCGACCTCTCG 658
DB 1247 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1306
QY 659 ACAAGGCACTTCCCTTCATGANGCCCTCCCTCCATGCTCATCCCGCCCTCC 718
DB 1307 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1366
QY 719 AGTCATGCCGCGACGAGTTAATATCAGTACCTCCGGGCTATCTCACCACACTTCC 778
DB 1367 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1426
QY 779 ACCTCCGAGTTTAAATAGTTTCCA 802
DB 1427 YYYYYYYYGTACCAATTTCTTA 1450

RESULT 6
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
```



```

;
; TITLE OF INVENTION: NOVEL RUCS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,087
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-965-087-1
;
; Query Match 1.3%; Score 62.8; DB 3; Length 732;
; Best Local Similarity 50.3%; Pred. No. 5.4e-07;
; Matches 154; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
;
; QY 3133 ACCACATGACGGTTGGAATTCCTGGTGATGCTGTGTGTAATTTCTGACGAGCGTCC 3192
; DB 158 ACCATAATGACGGTTAGAGTTTTCGGTGATGCGGTATTAGAAATTCACGGTTTCAGAT 217
;
; QY 3193 ATTTGTACTATTGTTTCCTAGTCTGGAAGAGGAGATTAGCAACCTATCGGACTGCCA 3252
; DB 218 ATTTATTTGATAACATCCCACTTGGCAGAGGGAATTTTAAACAAAATCGTGCCACTA 277
;
; QY 3253 TTGTTTCTGAGATCAGCACCTTGCATGCTAGCAAGAAACCTTGAACCTGGATCCATTTATGC 3312
; DB 278 TTGTATGTGAGCCCTCACTTGTAAATATTGCGAATAAATTTGGAATTTGAACGAATGATT 337
;
; QY 3313 TGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACCTTCGACATGCAATGGCCAAAT 3372
; DB 338 TACTTGGTAAAGGTGAAGAGAAACAGGGGACGCTACAGACCATCATTAATATCAGATG 397
;
; QY 3373 GTTTTGAAGCGTTAATAGGAGCTGTTTACTTGGAGGAGGAGCCCTGGAGGAGCCAGCACT 3432
; DB 398 CATTCGACATTTTATTTGGGCAATTGTTATTTGGATCAAGGACTAGATATAGTTTGGAAAT 457
;
; QY 3433 TATTTTG 3438
; DB 458 TTGCTG 463
;
; RESULT 9
; US-08-956-171E-25
; Sequence 25, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
;
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
;
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6251 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
;
; US-08-956-171E-25
;
; Query Match 1.3%; Score 62.4; DB 3; Length 6251;
; Best Local Similarity 50.0%; Pred. No. 3e-06;
; Matches 153; Conservative 1; Mismatches 152; Indels 0; Gaps 0;
;
; QY 3133 ACCACATGACGGTTGGAATTCCTGGTGATGCTGTGTGTAATTTCTGACGAGCGTCC 3192
; DB 5788 ACCATAATGACGGTTTAGAGTTTTCGGTGATGCGGTATTAGAAATTCACGGTTTCAGAT 5847
;
; QY 3193 ATTTGTACTATTGTTTTCCTAGTCTGGAAGAGGAGATTAGCAACCTATCGGACTGCCA 3252
; DB 5848 ATTTATTTGATAACATCCCACTTGGCAGAGGGAATTTTAAACAAAATCGTGCCACTA 5907
;
; QY 3253 TTGTTTCTGAGATCAGCACCTTGCATGCTAGCAAGAAACCTTGAACCTGGATCCATTTATGC 3312
; DB 5908 TTGTATGTGAGCCCTCACTKTGTAATATTTCGGAATAAATTTGGAATTTGAACGAATGATT 5967
;
; QY 3313 TGTATGCTCAGGGCCCTGACCTTTGTAGAGAAATCGGACCTTCGACATGCAATGGCCAAAT 3372
; DB 5968 TACTTGGTAAAGGTGAAGAGAAACAGGGGACGCTACAGACCATCATTAATATCAGATG 6027
;
; QY 3373 GTTTTGAAGCGTTAATAGGAGCTGTTTACTTGGAGGAGGAGCCCTGGAGGAGCCAGCACT 3432
; DB 6028 CATTCGACATTTTATTTGGGCAATTGTTATTTGGATCAAGGACTAGATATAGTTTGGAAAT 6087
;
; QY 3433 TATTTTG 3438
; DB 6088 TTGCTG 6093
;
; RESULT 10
```

us-10-774-974-1.rnd

Wed Dec 28 10:04:59 2005

Patent No. 6197510
 GENERAL INFORMATION:
 APPLICANT: Vinavagamoorthy, Thuraiayah
 TITLE OF INVENTION: Multi-Loci Genomic Analysis
 FILE REFERENCE: 44747
 CURRENT APPLICATION NUMBER: US/09/165,264
 CURRENT FILING DATE: 1998-10-01
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 13
 LENGTH: 320
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
 US-09-165-264-13

Query Match 1.3%; Score 61.6; DB 3; Length 320;
 Best Local Similarity 50.0%; Pred. No. 6.8e-07;
 Matches 154; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
 408 CCTCAAGTGCCTTCCACACACTTCTCAAACTCTCAGCCGCCAAATTTCTCCCTCCA 467
 309 CGTCCAGTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 250
 468 CGACAGACTTTGTACCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 527
 249 CC 190
 528 CCTGCCCCAATCAGGCGCTTTTCCCAACCAACAGATGAGGACGCTTCCCTCCCTCCCT 587
 189 CC 130
 588 CTTGTTTCTTCCCTCCATGCCACCAATGCTTGTCTTAATAACCCCCCAGTCCCTGGG 647
 129 CC 70
 648 GCACCTCTCTGGACAGGACACTTCCCTTTCATGATGCCCCCTCCCTCCCTCCCTCCCT 707
 69 CC 10
 708 CCGCCCCC 715
 9 CCCCCCCC 2

RESULT 12
 US-09-583-110-1893
 Sequence 1893, Application US/09583110
 Patent No. 6699703
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al.
 TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
 TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
 FILE REFERENCE: PATH00-07A
 CURRENT APPLICATION NUMBER: US/09/583,110
 CURRENT FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/107,433
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US 60/085,131
 PRIOR FILING DATE: 1998-05-12
 PRIOR APPLICATION NUMBER: US 60/051,553
 PRIOR FILING DATE: 1997-07-02
 NUMBER OF SEQ ID NOS: 5322
 SEQ ID NO 1893
 LENGTH: 699
 TYPE: DNA
 ORGANISM: Streptococcus pneumoniae
 US-09-583-110-1893
 Query Match 1.3%; Score 61; DB 3; Length 699;
 Best Local Similarity 49.2%; Pred. No. 1.7e-06;
 Matches 160; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

US-08-781-986A-25
 Sequence 25, Application US/08781986A
 Patent No. 6737248
 GENERAL INFORMATION:
 APPLICANT: Charles Kunech
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5255
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,986A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Benson, Bob
 REGISTRATION NUMBER: 30,446
 REFERENCE/DOCKET NUMBER: PB248PP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6251 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-781-986A-25

Query Match 1.3%; Score 62.4; DB 3; Length 6251;
 Best Local Similarity 50.0%; Pred. No. 3e-06;
 Matches 153; Conservative 1; Mismatches 152; Indels 0; Gaps 0;
 3133 ACCAATCAAGCGTTGGAATTCCTGGGTGATGCTGTTGTAATTTCTGACCAAGCTCC 3192
 5788 ACCATAATGAGCTTTAGATTTTGGGTGATGCGGTATTAGAAATGACGGTTTCCAGAT 5847
 3193 ATTGTGACTATTGTTTCTAGTCTGGAGAGGAGGATTAGCAACTATCGGACTGCCA 3252
 5848 ATTATTGTAACATCCCACTTCCCAAGAGGGAATTTACAAAATGCGTGCCACTA 5907
 3253 TTGTTCAAGATGACACCTTGCATGCTAGCAAGAACTTGAATCGGATCCATTTATGC 3312
 5908 TTGTATGAGCCCTCACTGTAATATTGGGATATAAATTGGATTGAACGAATGATTT 5967
 3313 TGATGCTCAGCGCCCTGACCTTTGTAGAGAAATCGGACCTTCGCATGCAATGCCAATT 3372
 6027 TACTTGGTAAGGTGAAGAGAAACAGGGGACGTACAGACCAATCATTAATATCAGATG 6072
 3372 GTTTGAAGCGTTAATAGAGCTGTTTACTTGGAGGAGCTGGAGGAAGCCAGCAAGT 3432
 6028 CATTCAAGCATTTATTGGGCAATTGTTTGGATCAAGGACTAGATATAGTTTGGAAAT 6087
 3433 TATTTG 3438
 6088 TTGCTG 6093
 RESULT 11
 US-09-165-264-13/c
 Sequence 13, Application US/09165264

us-10-774-974-1.rni

Wed Dec 28 10:04:59 2005

```

Db      208 ATGAGCGGATGGTGGTCAACCTTAGTGCAGGTATACCTCGCTGAACTAGCAGGTGAA 267
QY      3291 CTTGAACTGGATCATTTATGCTGTATGCTACAGGCGCTGACCTTTGTGTAGAGAAATCGGAC 3350
Db      268 TTTGAACTGGTGAATGTTTACGGTTAGGCGCGGAGAAATTAAGAGCGGTGTTTCGT 327
QY      3351 CTTGACATGCAATGGCAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTGA 3406
Db      328 CGCGAATCTATTTGGCCGACAGTAGAAGCGTTAATGTTGCTATTTTCTTGA 383

```

```

RESULT 15
US-09-165-264-11/c
; Sequence 11, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-11

```

```

Query Match      1.3%; Score 59.8; DB 3; Length 320;
Best Local Similarity 49.8%; Pred. No. 2.2e-06;
Matches 151; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY      416 TGCCCTTCCACCACTTCTCAAACTCTCCAGCGCCCAATTTCTCCCTCCACGACGAGA 475
Db      303 TGCCCTTCCACCACTTCTCAAACTCTCCAGCGCCCAATTTCTCCCTCCACGACGAGA 475
QY      476 CTTTGTACCTTTCCCTCCACCACTCTCGTCAAGCAAGGCGCTTTCCCTCCCTGCCC 535
Db      243 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 184
QY      536 AATCAGGCGCGCTTTCCCAACCAACAGATGAGGCAACCTTTCCAGTTCTCTGTTT 595
Db      183 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 124
QY      596 TCCTCCCTCCACCACTTCTCAAACTCTCCAGCGCCCAATTTCTCCCTCCACGACGAGA 655
Db      123 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 64
QY      656 TGGACAGGCACTTTCCCTCCACCACTCTCGTCAAGCAAGGCGCTTTCCAGTTCTCTGTT 715
Db      63 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 4
QY      716 TCC 718
Db      3 CCC 1

```

Search completed: December 27, 2005, 03:24:37
Job time : 782 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 15:11:52 ; Search time 3504 Seconds

(without alignments)
11242.956 Million cell updates/sec

Title: US-10-774-974-1

Perfect score: 4764

Sequence: 1 cgtcttgtagctcggtgta.....aaaaaaaaaaaaaaaaaaaa 4764

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.Main:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4764	100.0	4764	5 US-10-079-185-1	Sequence 1, Appli
2	4764	100.0	4764	7 US-10-774-974-1	Sequence 1, Appli
3	4764	100.0	4764	10 US-11-001-993-1	Sequence 1, Appli
4	4760.8	99.9	4764	3 US-09-900-425A-1	Sequence 1, Appli
5	4760.8	99.9	4764	8 US-10-805-919-1	Sequence 1, Appli
6	4760.8	99.9	4764	8 US-10-723-860-280	Sequence 280, App
7	4760.8	99.9	4764	8 US-10-756-149-319	Sequence 319, App
8	4732.4	99.3	5425	8 US-10-723-860-5003	Sequence 5003, App
9	1619.6	34.0	1626	7 US-10-205-331-117	Sequence 117, App
10	1398.6	29.4	1458	5 US-10-103-313-23	Sequence 187, App
11	1262.4	26.5	1314	5 US-10-103-313-187	Sequence 187, App
12	770.2	16.2	4275	10 US-11-097-143-2408	Sequence 2408, App
13	703.8	14.8	6397	10 US-11-097-143-2407	Sequence 2407, App
14	518.2	10.9	578	9 US-10-779-543-6006	Sequence 6006, App
15	469.4	9.9	633	8 US-10-357-930-20941	Sequence 20941, A
16	469.4	9.9	633	8 US-10-357-930-20985	Sequence 20985, A
17	469.4	9.9	633	8 US-10-357-930-26789	Sequence 26789, A
18	469.4	9.9	633	8 US-10-357-930-26829	Sequence 26829, A
19	411.4	8.6	488	3 US-09-918-995-20306	Sequence 20306, A
20	405.4	8.5	583	3 US-09-814-353-19214	Sequence 19214, A
21	392.2	8.2	405	3 US-09-803-713-409	Sequence 409, App
22	392.2	8.2	405	9 US-10-779-543-13679	Sequence 13679, A
23	384	8.1	397	9 US-10-779-543-9312	Sequence 9312, App

ALIGNMENTS

RESULT 1

US-10-079-185-1
; Sequence 1, Application US/10079185
; Publication No. US20030044941A1
; GENERAL INFORMATION:

; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS5030
; CURRENT APPLICATION NUMBER: US/10/079,185
; CURRENT FILING DATE: 2002-02-20

; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06

; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 4764

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-079-185-1

Query Match 100.0%; Score 4764; DB 5; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGCTCTGTGCTACTCGGCTAGTAGCTGCTTTGCTCTGACGCGCATCTCGGCGCCGAG	60
Db	1	CGTCTTGCTACTCGGCTAGTAGCTGCTTTGCTCTGACGCGCATCTCGGCGCCGAG	60
Qy	61	AGCCTTTTATAGTTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	120
Db	61	AGCCTTTTATAGTTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	120
Qy	121	CCCATATCATCAAGAGCTGATAAATCTAGTGAAGATTAGAGCTGTGCATCTTCACTA	180
Db	121	CCCATATCATCAAGAGCTGATAAATCTAGTGAAGATTAGAGCTGTGCATCTTCACTA	180
Qy	181	TGATATGAGGCAGTCTCTGAGCTTATATCTCTGGAAGATGTGACATATCCAGCGGA	240
Db	181	TGATATGAGGCAGTCTCTGAGCTTATATCTCTGGAAGATGTGACATATCCAGCGGA	240

Sequence 12001, A
Sequence 10556, A
Sequence 26, Appl
Sequence 10660, A
Sequence 10542, A
Sequence 2030, Ap
Sequence 31853, A
Sequence 40819, A
Sequence 40890, A
Sequence 12974, A
Sequence 1491, Ap
Sequence 199, App
Sequence 6589, Ap
Sequence 10743, A
Sequence 457, App
Sequence 295, App
Sequence 79823, A
Sequence 79824, A
Sequence 1050, Ap
Sequence 71764, A

24 381.8 8.0 394 9 US-10-779-543-12001
25 347.2 7.3 360 9 US-10-779-543-10556
c 26 329.8 6.9 342 3 US-09-815-343-26
c 27 329.8 6.9 342 7 US-10-097-105-26
28 319.4 6.7 465 8 US-10-357-930-10660
29 315 6.6 384 9 US-10-779-543-10542
30 300 6.3 300 9 US-10-779-543-2030
31 289.8 6.1 472 8 US-10-357-930-31853
32 289.8 6.1 472 8 US-10-357-930-40819
33 289.8 6.1 472 8 US-10-357-930-40890
34 223 4.7 336 3 US-09-814-353-12974
35 207.6 4.4 432 8 US-10-357-930-1491
36 206 4.3 326 3 US-09-814-353-199
37 206 4.3 326 3 US-09-814-353-6589
38 178.6 3.7 349 8 US-10-357-930-10743
c 39 174.4 3.7 263 7 US-10-469-285-457
40 169.4 3.6 464 3 US-09-728-445-295
41 169.4 3.6 464 9 US-10-964-549-295
42 161.6 3.4 2258 4 US-09-925-065A-79823
43 161.6 3.4 2258 4 US-09-925-065A-79824
44 112.4 2.4 470 3 US-09-864-761-1050
45 109.6 2.3 1729 4 US-09-925-065A-71764

QY 241 ACATCATGATGACGGGAACACATGTCACAGAACTGCTGCCACCGGGAGAGGGCGTC 300
DB 241 ACATCATGATGACGGGAACACATGTCACAGAACTGCTGCCACCGGGAGAGGGCGTC 300
QY 301 CCCGAGGACGAGGAGACATGAGCCAGACCCTCAGCACCACTCTTTAGGGCCCCAAAATC 360
DB 301 CCCGAGGACGAGGAGACATGAGCCAGACCCTCAGCACCACTCTTTAGGGCCCCAAAATC 360
QY 361 TGAGGCTGCTTCCACCCCTCAGACGCTCTCTGTGCAATATCAATATGAACTCCAAAGTGCCC 420
DB 361 TGAGGCTGCTTCCACCCCTCAGACGCTCTCTGTGCAATATCAATATGAACTCCAAAGTGCCC 420
QY 421 CTTTCAACAATCTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACACAGACTTTG 480
DB 421 CTTTCAACAATCTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACACAGACTTTG 480
QY 481 TACCCCTTCCCCCAACCACTCGCTCAGGCGCAAGGCCCTCTTCCCTCCCTGCCAAATCA 540
DB 481 TACCCCTTCCCCCAACCACTCGCTCAGGCGCAAGGCCCTCTTCCCTCCCTGCCAAATCA 540
QY 541 GGCAGGCTTTCGCCCAACCACTCAGATGAGGCAACCCCTTCCAGTTCCTCTTGTTCCTC 600
DB 541 GGCAGGCTTTCGCCCAACCACTCAGATGAGGCAACCCCTTCCAGTTCCTCTTGTTCCTC 600
QY 601 CCATGCAACAACAATGCTTGTCTTAATAAACCCGCCAGTCCCTGGGGCACCTCTGGAC 660
DB 601 CCATGCAACAACAATGCTTGTCTTAATAAACCCGCCAGTCCCTGGGGCACCTCTGGAC 660
QY 661 AAGGCACTTTCCTTTCATGATGCCCCCTCCCTCCATGCTCATCCCCCGCCCCCTCCAG 720
DB 661 AAGGCACTTTCCTTTCATGATGCCCCCTCCCTCCATGCTCATCCCCCGCCCCCTCCAG 720
QY 721 TCATGCGGAGCAGAGTTAATATACGTACCTCCGGGCTATTCTCACCAAACTTCCCAAC 780
DB 721 TCATGCGGAGCAGAGTTAATATACGTACCTCCGGGCTATTCTCACCAAACTTCCCAAC 780
QY 781 CTCGCCAGTTTAAATAGTTTCCAGAAACAAACCTAGTTCTTCTTCCCTCCAGTGCTAATAACA 840
DB 781 CTCGCCAGTTTAAATAGTTTCCAGAAACAAACCTAGTTCTTCTTCCCTCCAGTGCTAATAACA 840
QY 841 GCAGTAGTCTCATTTTCAGACATCTCCCTCATACCCACTCCCAAAAGGCTCCCAAGTGAGA 900
DB 841 GCAGTAGTCTCATTTTCAGACATCTCCCTCATACCCACTCCCAAAAGGCTCCCAAGTGAGA 900
QY 901 GAAGGTCCCAGAAAGGCTGAACACTATGATGACCAACAGGCAACCGAGACCAAGTCAAG 960
DB 901 GAAGGTCCCAGAAAGGCTGAACACTATGATGACCAACAGGCAACCGAGACCAAGTCAAG 960
QY 961 GCGAGGTGAGAGGCAATCGGTCCCTGGATCGCGGGAGCGAGCGCGAGTCCCGACAGGA 1020
DB 961 GCGAGGTGAGAGGCAATCGGTCCCTGGATCGCGGGAGCGAGCGCGAGTCCCGACAGGA 1020
QY 1021 GAAGACAGACAGCGGTACAGATCTGATTTATGACGAGGGAGAACACCACTCTCGCCACC 1080
DB 1021 GAAGACAGACAGCGGTACAGATCTGATTTATGACGAGGGAGAACACCACTCTCGCCACC 1080
QY 1081 GCAGCTACGAAACGAGCAGAGAGCGAGAACCGGAGAGACACAGGCATCGAGACAAACCGAA 1140
DB 1081 GCAGCTACGAAACGAGCAGAGAGCGAGAACCGGAGAGACACAGGCATCGAGACAAACCGAA 1140
QY 1141 GATCACCATCTCTGGAAAGGTCTTACAAAAGAGTATAAGATCTTGGAAAGGATTACG 1200
DB 1141 GATCACCATCTCTGGAAAGGTCTTACAAAAGAGTATAAGATCTTGGAAAGGATTACG 1200
QY 1201 GTTTATCGGTGTTCTTCAACCTGCTGATGTCACACAGAAATACCTGGGGAGATTATTA 1260
DB 1201 GTTTATCGGTGTTCTTCAACCTGCTGATGTCACACAGAAATACCTGGGGAGATTATTA 1260
QY 1261 AAAATACAGATTCTTGGGCCCCACCTCGAGATTGTGAATCATCGCTCCCCCAAGTAGGG 1320
DB 1261 AAAATACAGATTCTTGGGCCCCACCTCGAGATTGTGAATCATCGCTCCCCCAAGTAGGG 1320
QY 1321 AGAAGAGAGAGAGCTCGTTTGGGAGGAAGAAAGAACCGTTGGAGTGACAAACAGAGTTCTG 1380

DB 1321 AGAAGAGAGAGCTCGTTTGGGAGGAAGAAAGACCGTTGGAGTGACAAACAGAGTTCTG 1380
QY 1381 GCAAGACAGAACTATATACCTCAATCAAGGAAAGAGACCGGAGGAGACCATGCTGAC 1440
DB 1381 GCAAGACAGAACTATATACCTCAATCAAGGAAAGAGACCGGAGGAGACCATGCTGAC 1440
QY 1441 AGAATGAGGAGGAAGAAAGAACTTTCTTAAGCCCTGTGTGGATTCGATGCACTCATTCAG 1500
DB 1441 AGAATGAGGAGGAAGAAAGAACTTTCTTAAGCCCTGTGTGGATTCGATGCACTCATTCAG 1500
QY 1501 AAAAATACTACTCTCCAGTGAACCCATGATCGAGTGGGAGATTTCTACAGTGGTTGAAACGA 1560
DB 1501 AAAAATACTACTCTCCAGTGAACCCATGATCGAGTGGGAGATTTCTACAGTGGTTGAAACGA 1560
QY 1561 GTAGGCTTCTGTCATTTATATGACAAATTTGAGGAGGATTTGGGAGGAGGCAAGAAAGG 1620
DB 1561 GTAGGCTTCTGTCATTTATATGACAAATTTGAGGAGGATTTGGGAGGAGGCAAGAAAGG 1620
QY 1621 CCAGAGCTGCTCGGCTCCGTGGGAACCTCCAAAGCAAGCTCGATGAAGATTTTAGAGA 1680
DB 1621 CCAGAGCTGCTCGGCTCCGTGGGAACCTCCAAAGCAAGCTCGATGAAGATTTTAGAGA 1680
QY 1681 GTTCCAGTGAATCCGAGTGTGATCTGATGAGGACAGCACTGTCTAGCAGCTCAGACT 1740
DB 1681 GTTCCAGTGAATCCGAGTGTGATCTGATGAGGACAGCACTGTCTAGCAGCTCAGACT 1740
QY 1741 CTGAAGTTTTTGAAGTTTTGACAGAAATCAAAACGCAAAAGGCCCCCTGACCCGACTTC 1800
DB 1741 CTGAAGTTTTTGAAGTTTTTGAAGAAATCAAAACGCAAAAGGCCCCCTGACCCGACTTC 1800
QY 1801 ATGATGAACTTTGGTATCAACGATCCAGGCAAGATGATGACCACTCTGCAATGCA 1860
DB 1801 ATGATGAACTTTGGTATCAACGATCCAGGCAAGATGATGACCACTCTGCAATGCA 1860
QY 1861 GGCAGAAAGCAAGACGCAAGGAAATTAGGCAAGCAATTTATCTGAGAGAGGCCATCA 1920
DB 1861 GGCAGAAAGCAAGACGCAAGGAAATTAGGCAAGCAATTTATCTGAGAGAGGCCATCA 1920
QY 1921 AGCCCTGCTGCTCTATGACCAACATGCTGGGAGACTTTTCCACTACCCGATCAGAGTCT 1980
DB 1921 AGCCCTGCTGCTCTATGACCAACATGCTGGGAGACTTTTCCACTACCCGATCAGAGTCT 1980
QY 1981 CCCCCTACGAACTTTTAACTGACAGGCAACTGATAGAAATACGATGATCAGAGT 2040
DB 1981 CCCCCTACGAACTTTTAACTGACAGGCAACTGATAGAAATACGATGATCAGAGT 2040
QY 2041 ATATCTTTGAAGGATTTCTATGTTTGCATGCCCCCTGACCAATATTTCCACTGTGTA 2100
DB 2041 ATATCTTTGAAGGATTTCTATGTTTGCATGCCCCCTGACCAATATTTCCACTGTGTA 2100
QY 2101 AGTAATATGATTCACATAGACTACAGATTCATTTTCAATGAGAGATGATCCGAGGA 2160
DB 2101 AGTAATATGATTCACATAGACTACAGATTCATTTCAATGAGAGATGATCCGAGGA 2160
QY 2161 ATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTCTTATTCAGAGATATTTTGAAT 2220
DB 2161 ATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTCTTATTCAGAGATATTTTGAAT 2220
QY 2221 TATATGACTGGAATCTTAAAGGTCTTTTGTGGAAGACAGCCCTCTGCTGCTGCCAAGAT 2280
DB 2221 TATATGACTGGAATCTTAAAGGTCTTTTGTGGAAGACAGCCCTCTGCTGCTGCCAAGAT 2280
QY 2281 TTCATTTCAATGCGCAAGTTTGTGAAGATTTCTTCAGATGAGGAAAGAGTGTGTCGA 2340
DB 2281 TTCATTTCAATGCGCAAGTTTGTGAAGATTTCTTCAGATGAGGAAAGAGTGTGTCGA 2340
QY 2341 TGCAACAGATTTCTCTGATCTTGTAAAGGTGCAAGAACCCCTGGTCCCTGCTGAGGAGGA 2400
DB 2341 TGCAACAGATTTCTCTGATCTTGTAAAGGTGCAAGAACCCCTGGTCCCTGCTGAGGAGGA 2400
QY 2401 TTGCAATATGCTTTCAGTGGGAGGAGCTGGAGTGCGCAAGAAATATGCAAGAGATTCGAAG 2460

Db 2401 TTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATCCAGAAAGTCAAG 2460
QY 2461 GCATGATGTTTACCAACCTCGGAGCAAAACCAAGCTCTGTCCGTATCGATCAACATGGATC 2520
Db 2461 GCATGATGTTTACCAACCTCGGAGCAAAACCAAGCTCTGTCCGTATCGATCAACATGGATC 2520
QY 2521 GTGAACAGTTTCAACCCCGATGTGATTAATTCTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580
Db 2521 GTGAACAGTTTCAACCCCGATGTGATTAATTCTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580
QY 2581 CTGCACAGTTGAGTTATGTCAGGAGACCCACAGTACCAGAACTCTGGAAGAGTTATGTGA 2640
Db 2581 CTGCACAGTTGAGTTATGTCAGGAGACCCACAGTACCAGAACTCTGGAAGAGTTATGTGA 2640
QY 2641 AACTTCGCCACCTCTAGCAAAATAGTCCAAAGTCAAACTGA 2700
Db 2641 AACTTCGCCACCTCTAGCAAAATAGTCCAAAGTCAAACTGA 2700
QY 2701 CACAGAGGAGGAGCCCTCCAAAATAACGCGAGAGATACAAATGAGACGAGAGTAA 2760
Db 2701 CACAGAGGAGGAGCCCTCCAAAATAACGCGAGAGATACAAATGAGACGAGAGTAA 2760
QY 2761 CGGTGGAGCTAAAGTAGCAAGGATTTCTGGAAAACTGGCATCCGTTCTGTGATGTTCTGCAGC 2820
Db 2761 CGGTGGAGCTAAAGTAGCAAGGATTTCTGGAAAACTGGCATCCGTTCTGTGATGTTCTGCAGC 2820
QY 2821 ATGCAATGATGTACCTGTTCTGACCAATCATATPCGCTACCAACATGCTTAATGCATT 2880
Db 2821 ATGCAATGATGTACCTGTTCTGACCAATCATATPCGCTACCAACATGCTTAATGCATT 2880
QY 2881 TGGACAGTTGATAGGATATACTTTCAAGATCGTTGTCTGTGACGCTGGCCATGATCTC 2940
Db 2881 TGGACAGTTGATAGGATATACTTTCAAGATCGTTGTCTGTGACGCTGGCCATGATCTC 2940
QY 2941 ATCCAAAGTCAATTTAAATTTTGGAAATGAATCTCTGATCATGTCAGGAATTCATTAATCTA 3000
Db 2941 ATCCAAAGTCAATTTAAATTTTGGAAATGAATCTCTGATCATGTCAGGAATTCATTAATCTA 3000
QY 3001 ACTGTGAATTCGGCAGCCCAATACGGAGACAGAAAGTTTCATCATGCAATGCGGA 3060
Db 3001 ACTGTGAATTCGGCAGCCCAATACGGAGACAGAAAGTTTCATCATGCAATGCGGA 3060
QY 3061 AGAAGGATTAACACTTGATTAATATCATGTCACGCTTGGCCAGATGACCCAACTC 3120
Db 3061 AGAAGGATTAACACTTGATTAATATCATGTCACGCTTGGCCAGATGACCCAACTC 3120
QY 3121 CCTCGAGGATTAACCAATGAACGTTGGAAATTCCTGGTGTGATGCTGTTGTAATTTTC 3180
Db 3121 CCTCGAGGATTAACCAATGAACGTTGGAAATTCCTGGTGTGATGCTGTTGTAATTTTC 3180
QY 3181 TGACCAAGCTTCAATTTGTACTATTGTTTCTTAGTCTGGAAGAGGAGGATTAGCAACCT 3240
Db 3181 TGACCAAGCTTCAATTTGTACTATTGTTTCTTAGTCTGGAAGAGGAGGATTAGCAACCT 3240
QY 3241 ATCGGACTGCCATTGTTTCAAGATCAGACCTTGCCATGTGTAGCAAGAACTTTGAATCTGG 3300
Db 3241 ATCGGACTGCCATTGTTTCAAGATCAGACCTTGCCATGTGTAGCAAGAACTTTGAATCTGG 3300
QY 3301 ATCCATTTATGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAACTCGGACCTTGGACATG 3360
Db 3301 ATCCATTTATGCTGTATGCTCAGGGCCCTGACCTTTGTAGAGAACTCGGACCTTGGACATG 3360
QY 3361 CAATGGCCAAATTTGTTGAAGCGTTTAAATAGGAGCTGTTTACTTGGAGGAAAGCCTGGAGG 3420
Db 3361 CAATGGCCAAATTTGTTGAAGCGTTTAAATAGGAGCTGTTTACTTGGAGGAAAGCCTGGAGG 3420
QY 3421 AAGCCAAAGCAGTTAATTTGACGCTTGTCTTTAAATGATCCGACCTGCGCGAAGTCTGGC 3480
Db 3421 AAGCCAAAGCAGTTAATTTGACGCTTGTCTTTAAATGATCCGACCTGCGCGAAGTCTGGC 3480
QY 3481 TCAAATATCTCTCCACCCACTCCAACTACAGAGCCAAATACTGATCGACAACTTATTG 3540
Db 3481 TCAAATATCTCTCCACCCACTCCAACTACAGAGCCAAATACTGATCGACAACTTATTG 3540

QY 3541 AAACTTCTCCAGTTCTTACAAAACTTACTGAGTTTGAAGAGCAATTCGAGTAATTTTTTA 3600
Db 3541 AAACTTCTCCAGTTCTTACAAAACTTACTGAGTTTGAAGAGCAATTCGAGTAATTTTTTA 3600
QY 3601 CTCAATGTTGCACTTCTGGCAAGGCAATTCATTTGAGAACTGTGGGATTTAAACCATCTGA 3660
Db 3601 CTCAATGTTGCACTTCTGGCAAGGCAATTCATTTGAGAACTGTGGGATTTAAACCATCTGA 3660
QY 3661 CCTTAGGCCCAATCAGAGAAATGAAATTCCTAGGTGACTCCATTAATGCAATCTGGTAGCCA 3720
Db 3661 CCTTAGGCCCAATCAGAGAAATGAAATTCCTAGGTGACTCCATTAATGCAATCTGGTAGCCA 3720
QY 3721 CAGAGTACTTATTTCATTTCCAGATCATCTGAAGGACACTTAACTTTGTTGGCAA 3780
Db 3721 CAGAGTACTTATTTCATTTCCAGATCATCTGAAGGACACTTAACTTTGTTGGCAA 3780
QY 3781 GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATGCAAGAGT 3840
Db 3781 GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATGCAAGAGT 3840
QY 3841 ACGCCATAACCAACGACCAAGAGGCTGTGGCGCTTCGCCACCAAGACCTTGGCGG 3900
Db 3841 ACGCCATAACCAACGACCAAGAGGCTGTGGCGCTTCGCCACCAAGACCTTGGCGG 3900
QY 3901 ACCTTTTGGAAATCAATTTATTGTCAGCGCTGTACACTGATAAGGATTTGGAATATGTTTCATA 3960
Db 3901 ACCTTTTGGAAATCAATTTATTGTCAGCGCTGTACACTGATAAGGATTTGGAATATGTTTCATA 3960
QY 3961 CTTTTCATGAATGTCGTTCTTTCCAGATGTAAGAAATTCATTTTGAATCAGGATTTGA 4020
Db 3961 CTTTTCATGAATGTCGTTCTTTCCAGATGTAAGAAATTCATTTTGAATCAGGATTTGA 4020
QY 4021 ATGACCCCAATCCAGCTTTCAGCAGTGTGCTTGAACACTTAGGACAGAGGAAAAGAGC 4080
Db 4021 ATGACCCCAATCCAGCTTTCAGCAGTGTGCTTGAACACTTAGGACAGAGGAAAAGAGC 4080
QY 4081 CAGACATTCCTCTGTACAGACTCTTCAGACACAGTGGGCCCATCCCATGCCCCAACCTTACA 4140
Db 4081 CAGACATTCCTCTGTACAGACTCTTCAGACACAGTGGGCCCATCCCATGCCCCAACCTTACA 4140
QY 4141 CTGTGCTGTTTATTATTCAAAGGAGAAAGATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
Db 4141 CTGTGCTGTTTATTATTCAAAGGAGAAAGATAGGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
QY 4201 AAGCGAAATGGGAGCAGCAATGGATGCGCTTGAATAATATTAATTTTCCCAATGCCCC 4260
Db 4201 AAGCGAAATGGGAGCAGCAATGGATGCGCTTGAATAATATTAATTTTCCCAATGCCCC 4260
QY 4261 ATCAGAAAGCGGTTTCATCGAAACGGAAGTACAGACAGAGTTTAAAGAAATGAGGTGGGAAA 4320
Db 4261 ATCAGAAAGCGGTTTCATCGAAACGGAAGTACAGACAGAGTTTAAAGAAATGAGGTGGGAAA 4320
QY 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGAGGAGGCAT 4380
Db 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGAGGAGGCAT 4380
QY 4381 GCAAGTGTGGAGTATTACTTCTGCTCAGTAACTGTGTAATTAATTTTCCCAATGCCCC 4440
Db 4381 GCAAGTGTGGAGTATTACTTCTGCTCAGTAACTGTGTAATTAATTTTCCCAATGCCCC 4440
QY 4441 AGTTTTTCTGTCAGACAATGAACGAAAGTGTGCTCATTGAAATAAAATACAGAGTCAAAATCG 4500
Db 4441 AGTTTTTCTGTCAGACAATGAACGAAAGTGTGCTCATTGAAATAAAATACAGAGTCAAAATCG 4500
QY 4501 CTATTTGTTGTTTAAATGATCTGTTTTAGCTGGATGGTCTTTTATTAACAAAGTATTAGATT 4560
Db 4501 CTATTTGTTGTTTAAATGATCTGTTTTAGCTGGATGGTCTTTTATTAACAAAGTATTAGATT 4560
QY 4561 TTTCTTCTATTAAACGGAACCTTGACTTTGGTGAATGTGCATTTACTTCTTTTATTG 4620
Db 4561 TTTCTTCTATTAAACGGAACCTTGACTTTGGTGAATGTGCATTTACTTCTTTTATTG 4620

QY	1621	CCAAAGCTGCTCGSCCTCCGTGGGAACTCCAAAGACGAAGCTCGATGAAGATTTCAGAA	1680
Db	1621	CCAAAGCTGCTCGSCCTCCGTGGGAACTCCAAAGACGAAGCTCGATGAAGATTTCAGAA	1680
QY	1681	GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACCTGTTCTTAGCAGCTCAGACT	1740
Db	1681	GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACCTGTTCTTAGCAGCTCAGACT	1740
QY	1741	CTGAAGTTTTGAGGTTATTGCGAGAAATCAACGCAAAAAGGCCACCTGACCGACTTC	1800
Db	1741	CTGAAGTTTTGAGGTTATTGCGAGAAATCAACGCAAAAAGGCCACCTGACCGACTTC	1800
QY	1801	ATGATGAACCTTTGTTGATCAACGATCCAGGCCAGATGAATGATGACCACTCTGCAAAATGCA	1860
Db	1801	ATGATGAACCTTTGTTGATCAACGATCCAGGCCAGATGAATGATGACCACTCTGCAAAATGCA	1860
QY	1861	GCACAAGGCAAGACGCAAGGAAATTAGGACAGCACTTTATCTGGAGAAAGAGGCCATCA	1920
Db	1861	GCACAAGGCAAGACGCAAGGAAATTAGGACAGCACTTTATCTGGAGAAAGAGGCCATCA	1920
QY	1921	AGCCCTGTCCTATGACCAACAACTGCTGGGAGACTTTTCCACTACCGGATCACAGTCT	1980
Db	1921	AGCCCTGTCCTATGACCAACAACTGCTGGGAGACTTTTCCACTACCGGATCACAGTCT	1980
QY	1981	CCCGGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCAGAGT	2040
Db	1981	CCCGGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCAGAGT	2040
QY	2041	ATATCTTTGAAGGATTTCTATGTTTGCATGCCCCCTCGACCAATATTCCACTGTGTA	2100
Db	2041	ATATCTTTGAAGGATTTCTATGTTTGCATGCCCCCTCGACCAATATTCCACTGTGTA	2100
QY	2101	AAGTAATTAGATTCACATAGACTACACGATTCATTTCATGGAAGAGATGATCCGGAGA	2160
Db	2101	AAGTAATTAGATTCACATAGACTACACGATTCATTTCATGGAAGAGATGATCCGGAGA	2160
QY	2161	ATTTTGTGTGAAGGCTTGAATCTTTTCACTGTTCTTCACTTTCAGAGATATTTTGGAT	2220
Db	2161	ATTTTGTGTGAAGGCTTGAATCTTTTCACTGTTCTTCACTTTCAGAGATATTTTGGAT	2220
QY	2221	TATATGACTGGAATCTTTAAAGTCTTTGTTGAAGACAGCCCTCCCTGCTGCCCAAGAT	2280
Db	2221	TATATGACTGGAATCTTTAAAGTCTTTGTTGAAGACAGCCCTCCCTGCTGCCCAAGAT	2280
QY	2281	TTCAATTTCAATGCAAGTTTGAAGATTTCTTCAGATGAGGAAAGAGTCTGTCCA	2340
Db	2281	TTCAATTTCAATGCAAGTTTGAAGATTTCTTCAGATGAGGAAAGAGTCTGTCCA	2340
QY	2341	TGCACAGATTTCTCTGTACTTGAAGTGAAGGACCAAGCCCTGCTGAGGAGAGA	2400
Db	2341	TGCACAGATTTCTCTGTACTTGAAGTGAAGGACCAAGCCCTGCTGAGGAGAGA	2400
QY	2401	TTGCCAATATGCTTCAGTGGAGAGCTGAGTGGCAGAAATATGAGAAAGATGCAAG	2460
Db	2401	TTGCCAATATGCTTCAGTGGAGAGCTGAGTGGCAGAAATATGAGAAAGATGCAAG	2460
QY	2461	GCAATGATTTGTTACCAACCCCTGGGACCAACCAAGCTCTGTCCGATCGATCACTGCATC	2520
Db	2461	GCAATGATTTGTTACCAACCCCTGGGACCAACCAAGCTCTGTCCGATCGATCACTGCATC	2520
QY	2521	GTGAACAGTTTCAACCCCGATGTGATTTATCTTTCCGATTTATCGTCCAATTTGGGATACGCC	2580
Db	2521	GTGAACAGTTTCAACCCCGATGTGATTTATCTTTCCGATTTATCGTCCAATTTGGGATACGCC	2580
QY	2581	CTGCAGATTGATGACGAGAGCCACAGTACCAAAAACCTGTGGAAGATTATGTGA	2640
Db	2581	CTGCAGATTGATGACGAGAGCCACAGTACCAAAAACCTGTGGAAGATTATGTGA	2640
QY	2641	AATCTCGCCACCTCTAGCAAAATAGTCCCAAGTCAAACTGACAAACAGAAAGCTGG	2700
Db	2641	AATCTCGCCACCTCTAGCAAAATAGTCCCAAGTCAAACTGACAAACAGAAAGCTGG	2700
QY	2701	CACAGAGGGAGGAGCCCTCCAAAAAATACGGCAGAAAGATACAAATGAGACGGAAGTAA	2760

Db	2701	CACAGAGGGAGGAGCCCTCCAAAAAATACGGCAGAAAGATACAAATGAGACGGAAGTAA	2760
QY	2761	CGGTGGAGCTTAAGTAGCAAGGATTTCTGGAAAACTGGCATCCGTTCTGTCTGTGCTGAGC	2820
Db	2761	CGGTGGAGCTTAAGTAGCAAGGATTTCTGGAAAACTGGCATCCGTTCTGTCTGTGCTGAGC	2820
QY	2821	ATGCATGATGCTACCTGTTCTGACCCATCATATCCCTGCTACCAATGCCTAATGCATT	2880
Db	2821	ATGCATGATGCTACCTGTTCTGACCCATCATATCCCTGCTACCAATGCCTAATGCATT	2880
QY	2881	TGACAAAGTTGATAGGATATCTTTTCCAAAGATCGTGTCTGTGCTGAGCTGGCCATGACTC	2940
Db	2881	TGACAAAGTTGATAGGATATCTTTTCCAAAGATCGTGTCTGTGCTGAGCTGGCCATGACTC	2940
QY	2941	ATCCAAAGTCAATTTAAATTTTGGAAATGAATCTGATCATGTCAGGAATTCATTATCTA	3000
Db	2941	ATCCAAAGTCAATTTAAATTTTGGAAATGAATCTGATCATGTCAGGAATTCATTATCTA	3000
QY	3001	ACTGTGGAATTCGGAGCCCAAAATACGGACACAGAAAGTTTCATCATGCAATGCGGA	3060
Db	3001	ACTGTGGAATTCGGAGCCCAAAATACGGACACAGAAAGTTTCATCATGCAATGCGGA	3060
QY	3061	AGAAAGGATTAACACCTTGATAAATATCATGTACAGCCCTTGGCCAAAGATGACCAACTC	3120
Db	3061	AGAAAGGATTAACACCTTGATAAATATCATGTACAGCCCTTGGCCAAAGATGACCAACTC	3120
QY	3121	CCTCGAGGATTAACCAATGAAACGGTTGGAAATCTCTGGGTGATGCTGTTGTAATTC	3180
Db	3121	CCTCGAGGATTAACCAATGAAACGGTTGGAAATCTCTGGGTGATGCTGTTGTAATTC	3180
QY	3181	TGACCAAGCTTCATTTGTACTATTTGTTCTCTAGTCTGGAAGAGGAGATTAGCAACTG	3240
Db	3181	TGACCAAGCTTCATTTGTACTATTTGTTCTCTAGTCTGGAAGAGGAGATTAGCAACTG	3240
QY	3241	ATCGGACTGCAATGTTTCAGAAATCAGCACTTCCCATGCTAGCAAGAAATCTGAACTGG	3300
Db	3241	ATCGGACTGCAATGTTTCAGAAATCAGCACTTCCCATGCTAGCAAGAAATCTGAACTGG	3300
QY	3301	ATCCATTTATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATG	3360
Db	3301	ATCCATTTATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATG	3360
QY	3361	CAATGCCAATTTGTTTGAAGCTTAAATAGGAGCTGTTTACTTGGAGGGAAGCTGAGG	3420
Db	3361	CAATGCCAATTTGTTTGAAGCTTAAATAGGAGCTGTTTACTTGGAGGGAAGCTGAGG	3420
QY	3421	AAGCCAAAGCAGTTTATTTGGAGCTTCTTTTAAATGATCGGACCTGCGGAAAGTCTGGC	3480
Db	3421	AAGCCAAAGCAGTTTATTTGGAGCTTCTTTTAAATGATCGGACCTGCGGAAAGTCTGGC	3480
QY	3481	TCAATTTATCTCTCCAACCACTCAAGAGCCAAATCTGATCGACAACTTATTG	3540
Db	3481	TCAATTTATCTCTCCAACCACTCAAGAGCCAAATCTGATCGACAACTTATTG	3540
QY	3541	AAACTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATGGAAGTAAATTTTA	3600
Db	3541	AAACTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATGGAAGTAAATTTTA	3600
QY	3601	CTCATGTTTCGACTTCTGGCAAGGGCAATTCATTGAGAACTGTGGGATTTAACTCTGA	3660
Db	3601	CTCATGTTTCGACTTCTGGCAAGGGCAATTCATTGAGAACTGTGGGATTTAACTCTGA	3660
QY	3661	CCTTAGGCCCAATCAGAGAAATGGAATCTCTAGGTGACCTCAATATGCAACTGATGACCA	3720
Db	3661	CCTTAGGCCCAATCAGAGAAATGGAATCTCTAGGTGACCTCAATATGCAACTGATGACCA	3720
QY	3721	CAGAGTACTTATTTCAATCTATTTCCAGATCATCAATGAAGGACACTTTAACTTTGTCGAA	3780
Db	3721	CAGAGTACTTATTTCAATCTATTTCCAGATCATCAATGAAGGACACTTTAACTTTGTCGAA	3780
QY	3781	GCCTTTTGTGTAATATAGAACTCAGGCCAAGGTAGCGGAGGCTGGGGATCAGAGT	3840

		TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof	
		FILE REFERENCE: ISIS0003-104 (ISIS-5030US.D1)	
		CURRENT APPLICATION NUMBER: US/11/001,993	
		PRIOR FILING DATE: 2004-12-02	
		PRIOR APPLICATION NUMBER: 10/079,185	
		PRIOR FILING DATE: 2002-02-20	
		PRIOR APPLICATION NUMBER: 09/479,783	
		PRIOR FILING DATE: 2000-01-07	
		PRIOR APPLICATION NUMBER: 08/870,608	
		PRIOR FILING DATE: 1997-06-06	
		PRIOR APPLICATION NUMBER: 08/659,440	
		PRIOR FILING DATE: 1996-06-06	
		PRIOR APPLICATION NUMBER: 09/900,425	
		PRIOR FILING DATE: 2001-07-06	
		NUMBER OF SEQ ID NOS: 36	
		SOFTWARE: Patentin version 3.1	
		SEQ ID NO 1	
		LENGTH: 4764	
		TYPE: DNA	
		ORGANISM: Homo sapiens	
		US-11-001-993-1	
		Query Match 100.0%; Score 4764; DB 10; Length 4764;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 4764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	3781	GCTCTTTGGTGAATATAGAACTCAGGCCAAGTAGCGAGGAGCTGGGCATCGAGGACT	3840
Qy	3841	AGCCATACCAACGACGACGACGAGGCGCTGTGGCGCTTCGCACCAAGACCTTTGGCGG	3900
Db	3841	AGCCATACCAACGACGACGAGGCGCTGTGGCGCTTCGCACCAAGACCTTTGGCGG	3900
Qy	3901	ACCTTTTGAATCATTTATTGTGAGCGCTGTACACTGATGAAGATTGGAAATATGTTCTA	3960
Db	3901	ACCTTTTGAATCATTTATTGTGAGCGCTGTACACTGATGAAGATTGGAAATATGTTCTA	3960
Qy	3961	CTTTGATGATGCTGCTCTTTCCACGATTCGAAAGAAATTCATTTTGAATCAGGATTGA	4020
Db	3961	CTTTGATGATGCTGCTCTTTCCACGATTCGAAAGAAATTCATTTTGAATCAGGATTGA	4020
Qy	4021	ATGACCCCAATCCACAGCTTCAGCAGTGTCTTGACACTTAGGACAGAGGAAAGAGC	4080
Db	4021	ATGACCCCAATCCACAGCTTCAGCAGTGTCTTGACACTTAGGACAGAGGAAAGAGC	4080
Qy	4081	CAGACATTCCTCTGTACAGACTCTCGACAGTGGGCCCATCCCATGCCCGAACCTACA	4140
Db	4081	CAGACATTCCTCTGTACAGACTCTCGACAGTGGGCCCATCCCATGCCCGAACCTACA	4140
Qy	4141	CTGTGCTGTTTATTTCAAGGAGAAAGATAGCTGTGGAAAGACCAAGTATTTCAGC	4200
Db	4141	CTGTGCTGTTTATTTCAAGGAGAAAGATAGCTGTGGAAAGACCAAGTATTTCAGC	4200
Qy	4201	AAGCGAAATGGGACGACGAAATGGATGGCTGTGAAAAATATATTTTCCCGATGGCCC	4260
Db	4201	AAGCGAAATGGGACGACGAAATGGATGGCTGTGAAAAATATATTTTCCCGATGGCCC	4260
Qy	4261	ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTTAAAGAAATGAGTGGGAA	4320
Db	4261	ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTTAAAGAAATGAGTGGGAA	4320
Qy	4321	GAGAGCATCAAGAGAGAGCCGATGAGACTGAGACATCAAGAAATTAAGAGGCGCAT	4380
Db	4321	GAGAGCATCAAGAGAGAGCCGATGAGACTGAGACATCAAGAAATTAAGAGGCGCAT	4380
Qy	4381	GCAAGTGTGAGTATTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT	4440
Db	4381	GCAAGTGTGAGTATTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT	4440
Qy	4441	AGTTTTCTCTGACGACAAATGAACGAAGTGTCTCAATGAAATAAAATACAGAGTCAATCG	4500
Db	4441	AGTTTTCTCTGACGACAAATGAACGAAGTGTCTCAATGAAATAAAATACAGAGTCAATCG	4500
Qy	4501	CTATTGTTGTTTAAATCATCTGTTTGTAGCTGGATGCTTTTATTAACAAGTATTAGATT	4560
Db	4501	CTATTGTTGTTTAAATCATCTGTTTGTAGCTGGATGCTTTTATTAACAAGTATTAGATT	4560
Qy	4561	TTTCTCTTATTTAACGGAAACTTGACTTTTGGTGAATGTGCATTACTTCTTTTATTTTTG	4620
Db	4561	TTTCTCTTATTTAACGGAAACTTGACTTTTGGTGAATGTGCATTACTTCTTTTATTTTTG	4620
Qy	4621	CTCTTTAAATAATAAATTCGAAGACATATTCATGTGGAATAGATCTGTTTTTCCAT	4680
Db	4621	CTCTTTAAATAATAAATTCGAAGACATATTCATGTGGAATAGATCTGTTTTTCCAT	4680
Qy	4681	CTGTGCTCCAGATTGTGACCTTAGACTTTCAATTGACAAAGTAAAAAATGACTTTACTAG	4740
Db	4681	CTGTGCTCCAGATTGTGACCTTAGACTTTCAATTGACAAAGTAAAAAATGACTTTACTAG	4740
Qy	4741	TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT	4764
Db	4741	TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT	4764

QY 721 TCATGCCGAGCAGGTTAATTATCAGTACCTCTCGGGCTATTCTCAGCACAACCTTCCCAC 780
DB 721 TCATGCCGAGCAGGTTAATTATCAGTACCTCTCGGGCTATTCTCAGCACAACCTTCCCAC 780
QY 781 CTCCTCAGTTTAAATAGTTTCCAGAAACAACCTTAGTTCTTTCTGCTCCAGTGTCTAATAACA 840
DB 781 CTCCTCAGTTTAAATAGTTTCCAGAAACAACCTTAGTTCTTTCTGCTCCAGTGTCTAATAACA 840
QY 841 GCAGTAGTCTCTCATTTTCCAGATCTCCCTCCATACCCACTCTCCANAGGCTCCAGTGAGA 900
DB 841 GCAGTAGTCTCTCATTTTCCAGATCTCCCTCCATACCCACTCTCCANAGGCTCCAGTGAGA 900
QY 901 GAAGTCTCCAGAAAGGCTGAACACTATGATGACACAGGCCAGGACCGAGACCAAGTCATG 960
DB 901 GAAGTCTCCAGAAAGGCTGAACACTATGATGACACAGGCCAGGACCGAGACCAAGTCATG 960
QY 961 GCGCAGGTGAGAGGCATCGGTCCCTCGATCGCGGAGCGAGCGCAGTCCCGACAGGA 1020
DB 961 GCGCAGGTGAGAGGCATCGGTCCCTCGATCGCGGAGCGAGCGCAGTCCCGACAGGA 1020
QY 1021 GAAGACAAGACAGCCGGTACAGATCTGATTTATGACCGAGGAGAAACACCATCTCGCCACC 1080
DB 1021 GAAGACAAGACAGCCGGTACAGATCTGATTTATGACCGAGGAGAAACACCATCTCGCCACC 1080
QY 1081 GCAGCTACGAACGGAGCAGAGCGAGAAACGGGAGAGACAGGCAATCGAGACAACCGAA 1140
DB 1081 GCAGCTACGAACGGAGCAGAGCGAGAAACGGGAGAGACAGGCAATCGAGACAACCGAA 1140
QY 1141 GATCACCATCTCTGGAAGGCTCTCAAAAAGAGTATTAAGAGATCTGGAAGGAGTTACG 1200
DB 1141 GATCACCATCTCTGGAAGGCTCTCAAAAAGAGTATTAAGAGATCTGGAAGGAGTTACG 1200
QY 1201 GTTTATCGTGTCTCTGAACTCTGATGACACACAGAAATTACTCTGGGAGATTTATTA 1260
DB 1201 GTTTATCGTGTCTCTGAACTCTGATGACACACAGAAATTACTCTGGGAGATTTATTA 1260
QY 1261 AAAATACAGATCTTTGGGCCCCCACCCTCTGGAGATTTGTAATCATCGCTCCCCAAGTAGGG 1320
DB 1261 AAAATACAGATCTTTGGGCCCCCACCCTCTGGAGATTTGTAATCATCGCTCCCCAAGTAGGG 1320
QY 1321 AGAAGACAGAGCTCTGTTGGGAGGAGAAAGACCGTTGGAGTACAAACAGAGTTCTG 1380
DB 1321 AGAAGACAGAGCTCTGTTGGGAGGAGAAAGACCGTTGGAGTACAAACAGAGTTCTG 1380
QY 1381 GCAAGACAAGAACTATACCTCAATCAAGGAAAAGAGCCCGAGGAGACCATGCTGACA 1440
DB 1381 GCAAGACAAGAACTATACCTCAATCAAGGAAAAGAGCCCGAGGAGACCATGCTGACA 1440
QY 1441 AGAATGAGGAGGAGAAAGAACTTTCTTAAGCCTGTGTGATTCGATTCGACTCAATTCAG 1500
DB 1441 AGAATGAGGAGGAGAAAGAACTTTCTTAAGCCTGTGTGATTCGATTCGACTCAATTCAG 1500
QY 1501 AAAAATACTACTCCAGTACACCCATGGATCAGTGGGAGATTTCTACAGTGTGGAGCGA 1560
DB 1501 AAAAATACTACTCCAGTACACCCATGGATCAGTGGGAGATTTCTACAGTGTGGAGCGA 1560
QY 1561 GTAGGCTTCGTGACTTATGACAAATTTGAGGAGGAGTTGCGGAGCAGGCAAGAAAGG 1620
DB 1561 GTAGGCTTCGTGACTTATGACAAATTTGAGGAGGAGTTGCGGAGCAGGCAAGAAAGG 1620
QY 1621 CCAAAGCTCTCGGCCCTCGTGGGAACCTCCAAAGACGAAGCTCGATGGAAGATTTAGAGA 1680
DB 1621 CCAAAGCTCTCGGCCCTCGTGGGAACCTCCAAAGACGAAGCTCGATGGAAGATTTAGAGA 1680
QY 1681 GTTCAGTGAATCCAGTGTGAGTCTGATGAGGACAGCACTCTGTTCTAGCAGCTCAGACT 1740
DB 1681 GTTCAGTGAATCCAGTGTGAGTCTGATGAGGACAGCACTCTGTTCTAGCAGCTCAGACT 1740
QY 1741 CTGAAGTTTTGAGCTTTATTCAGAAATCAAAACCAAAAAGGCCACCTGACCGACTTC 1800
DB 1741 CTGAAGTTTTGAGCTTTATTCAGAAATCAAAACCAAAAAGGCCACCTGACCGACTTC 1800

QY 1801 ATGATGAACCTTTGGTACAACGATCCAGGCCAGATGAATGATGGACCACTCTGCAAAATGCA 1860
DB 1801 ATGATGAACCTTTGGTACAACGATCCAGGCCAGATGAATGATGGACCACTCTGCAAAATGCA 1860
QY 1861 GCGCAAGGCAAGACGACAGGAAATTAGGACACAGCATTTATCTCTGAGAAAGAGGCCATCA 1920
DB 1861 GCGCAAGGCAAGACGACAGGAAATTAGGACACAGCATTTATCTCTGAGAAAGAGGCCATCA 1920
QY 1921 AGCCCTGTCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCACAAGTCT 1980
DB 1921 AGCCCTGTCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGGATCACAAGTCT 1980
QY 1981 CCCCCTACGAACTTTTAACTGACAGGCCAACTCTGTTATAGAAATACGATGATCAGAGT 2040
DB 1981 CCCCCTACGAACTTTTAACTGACAGGCCAACTCTGTTATAGAAATACGATGATCAGAGT 2040
QY 2041 ATATCTTTGAAGGATTTCTATGTTTGACATGCCCCCTGACCAATATTTCCACTGTGTA 2100
DB 2041 ATATCTTTGAAGGATTTCTATGTTTGACATGCCCCCTGACCAATATTTCCACTGTGTA 2100
QY 2101 AAGTAATTAGATTCAACATAGACTACAGATTCATTTCAATGAAGAGATGATGCCGAGA 2160
DB 2101 AAGTAATTAGATTCAACATAGACTACAGATTCATTTCAATGAAGAGATGATGCCGAGA 2160
QY 2161 ATTTTGTGTGAAGGGCTTGAATCTTTTCACTGTCTCTATTCAGAGATATTTTGGAAAT 2220
DB 2161 ATTTTGTGTGAAGGGCTTGAATCTTTTCACTGTCTCTATTCAGAGATATTTTGGAAAT 2220
QY 2221 TATATGACTGGANCTTTAAAGGTCCTTTGTTGAAGACAGCCCTCTCTGCTGCCCAAGAT 2280
DB 2221 TATATGACTGGANCTTTAAAGGTCCTTTGTTGAAGACAGCCCTCTCTGCTGCCCAAGAT 2280
QY 2281 TTCAATTCATGCGATTTGTAAGATTTCTTCAGATGGAAGAAAGGCTGTGCTCA 2340
DB 2281 TTCAATTCATGCGATTTGTAAGATTTCTTCAGATGGAAGAAAGGCTGTGCTCA 2340
QY 2341 TGCAACAGATTTCTCTGTACTTTGTTAAGGTGAGCAAAAGCCCTGCTGCTGAGGAGGAGA 2400
DB 2341 TGCAACAGATTTCTCTGTACTTTGTTAAGGTGAGCAAAAGCCCTGCTGCTGAGGAGGAGA 2400
QY 2401 TTGCCAATATCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAG 2460
DB 2401 TTGCCAATATCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAG 2460
QY 2461 GCATGATTTGTACCAACCCCTGGAGCAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
DB 2461 GCATGATTTGTACCAACCCCTGGAGCAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
QY 2521 GTGAACAGTTTCAACCCCGATGTGATTTCTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580
DB 2521 GTGAACAGTTTCAACCCCGATGTGATTTCTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580
QY 2581 CTGCACAGTTGAGTTATGCAAGGAGACCAAGTACCAAACTGTTGGAAAGATTTATGTA 2640
DB 2581 CTGCACAGTTGAGTTATGCAAGGAGACCAAGTACCAAACTGTTGGAAAGATTTATGTA 2640
QY 2641 AACTTCCGCTCTCTAGCAAAATAGTCCCAAGTCAACAACTGCAAAAAGCTGG 2700
DB 2641 AACTTCCGCTCTCTAGCAAAATAGTCCCAAGTCAACAACTGCAAAAAGCTGG 2700
QY 2701 CACAGAGGAGGAGGCCCTCCAAAATAACCGCAGAAAGAAATCAATGAGACGAGAGTAA 2760
DB 2701 CACAGAGGAGGAGGCCCTCCAAAATAACCGCAGAAAGAAATCAATGAGACGAGAGTAA 2760
QY 2761 CGGTGGAGCTTAAGTAGCAAGGATTTGGAAAACTGGCATCCGTTCTGATGCTGTGTCAGC 2820
DB 2761 CGGTGGAGCTTAAGTAGCAAGGATTTGGAAAACTGGCATCCGTTCTGATGCTGTGTCAGC 2820
QY 2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCTTAATGATT 2880
DB 2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCTTAATGATT 2880
QY 2881 TGGACAAGTTTGATAGGATATATCTTTTCCAAGATCGTTGCTGTGTCAGCTGGCCATGACTC 2940

2881 TGACAAAGTTGATAGGATATACATTTTCAAGATGCTGTTGCTGCTGAGCTGGCCATGACTC 2940
2941 ATCCAGTCATCATTTTAAATTTTGGAAATGAATCTGATCATGTCAGGAAATTCATTAATCTA 3000
2941 ATCCAGTCATCATTTTAAATTTTGGAAATGAATCTGATCATGTCAGGAAATTCATTAATCTA 3000
3001 ACTGTGAAATTCGGCAGCCCAATACCGAGACAGAAAGTTTCATCATGACATGCGGA 3060
3001 ACTGTGAAATTCGGCAGCCCAATACCGAGACAGAAAGTTTCATCATGACATGCGGA 3060
3061 AGAAAGGAGTTAACACCTTGATTAATATCATGTCAAGCTTTGGCCAGATGACCAACTC 3120
3061 AGAAAGGAGTTAACACCTTGATTAATATCATGTCAAGCTTTGGCCAGATGACCAACTC 3120
3121 CCTCAGGAGTTAACACCAATGAACCGTTTGGAAATTCCTGGGTGATGCTGTTGTAATTTTC 3180
3121 CCTCAGGAGTTAACACCAATGAACCGTTTGGAAATTCCTGGGTGATGCTGTTGTAATTTTC 3180
3181 TGACAGCGCTCCATTTGTATCTATTTGTTTCTAGTCTGGAAGAGGAGATTAGCAACT 3240
3181 TGACAGCGCTCCATTTGTATCTATTTGTTTCTAGTCTGGAAGAGGAGATTAGCAACT 3240
3241 ATCGGAGTGCATGTTTCAGAAATCAGACCTTGCCATGCTAGCAAAAGAACTTGAAGCTG 3300
3241 ATCGGAGTGCATGTTTCAGAAATCAGACCTTGCCATGCTAGCAAAAGAACTTGAAGCTG 3300
3301 ATCCATTTATGCTGATGCTCAGCGGCTGACCTTTGTAGAGAATCGGACCTTCGACATG 3360
3301 ATCCATTTATGCTGATGCTCAGCGGCTGACCTTTGTAGAGAATCGGACCTTCGACATG 3360
3361 CAATGCGCAATTTGTTTGAAGCGTTAATAGAGCTGTTTACTTGGAGGAGCGCTCGAGG 3420
3361 CAATGCGCAATTTGTTTGAAGCGTTAATAGAGCTGTTTACTTGGAGGAGCGCTCGAGG 3420
3421 AAGCCAGCAGTTATTTGGAAGCGTTTAAATGATTCGGACCTGCGGAGCTGCTGGC 3480
3421 AAGCCAGCAGTTATTTGGAAGCGTTTAAATGATTCGGACCTGCGGAGCTGCTGGC 3480
3481 TCATTTATCTCTCCACCCACTTCCAACTACAGAGCCAAATCTGATGACAACTTTATG 3540
3481 TCATTTATCTCTCCACCCACTTCCAACTACAGAGCCAAATCTGATGACAACTTTATG 3540
3541 AAACCTTCTCCAGTTCTACAAAATCTTACTGAGTTTGAAGCAATTTGAGTAATTTTGA 3600
3541 AAACCTTCTCCAGTTCTACAAAATCTTACTGAGTTTGAAGCAATTTGAGTAATTTTGA 3600
3601 CTGATGTTTGAATTTGAGGAGCTTCAATTTGAGAACTGTTGGGATTTAACCATCTGA 3660
3601 CTGATGTTTGAATTTGAGGAGCTTCAATTTGAGAACTGTTGGGATTTAACCATCTGA 3660
3661 CCTAGGCGCAATCAGAGATGAATTTCTTAGTGATCTCCATATGCAACTGAGTGGACCA 3720
3661 CCTAGGCGCAATCAGAGATGAATTTCTTAGTGATCTCCATATGCAACTGAGTGGACCA 3720
3721 CAGAGTACTTATTCATTTCCAGATGGAATTTCCAGATGGAATTTGAGTGGGATTTGCGAA 3780
3721 CAGAGTACTTATTCATTTCCAGATGGAATTTCCAGATGGAATTTGAGTGGGATTTGCGAA 3780
3721 CAGAGTACTTATTCATTTCCAGATGGAATTTCCAGATGGAATTTGAGTGGGATTTGCGAA 3780
3781 GCTCTTTGTTGTAATTAATGAACTCAGGCGCAAGTAGCGGAGCTGCGCAAGAGCTTGGCGG 3840
3781 GCTCTTTGTTGTAATTAATGAACTCAGGCGCAAGTAGCGGAGCTGCGCAAGAGCTTGGCGG 3840
3841 ACGCCATAACCAACGACAGACCAAGAGCGCTTGGCGCTTGGCACCAGAGCTTGGCGG 3900
3841 ACGCCATAACCAACGACAGACCAAGAGCGCTTGGCGCTTGGCACCAGAGCTTGGCGG 3900
3901 ACCTTTTGGAAATCAATTTATTTGACGCGCTGATGATGAGGATTTGGAAATGTTTCATA 3960
3901 ACCTTTTGGAAATCAATTTATTTGACGCGCTGATGATGAGGATTTGGAAATGTTTCATA 3960
3961 CTTTTCATGAATGCTGCTCTTTTCCAGATTTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020

3961 CTTTTCATGAATGCTGCTCTTTTCCAGATTTGAAAGAAATTCATTTTGAATCAGGATTTGA 4020
4021 ATGACCCCAATCCAGCTTTTCCAGAGTGTGTTGTTGACCTTAGGACAGAGGAAAGAGC 4080
4021 ATGACCCCAATCCAGCTTTTCCAGAGTGTGTTGTTGACCTTAGGACAGAGGAAAGAGC 4080
4081 CAGACATCTCTGTATCAAGAGCTCTGACAGAGTGGGCGCCATCCCATGCCCCGAACTTACA 4140
4081 CAGACATCTCTGTATCAAGAGCTCTGACAGAGTGGGCGCCATCCCATGCCCCGAACTTACA 4140
4141 CTGTGGCTGTTTATTTTCAAGGAGAGAAATAGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
4141 CTGTGGCTGTTTATTTTCAAGGAGAGAAATAGCTGTGGGAAAGGACCAAGTATTTCAGC 4200
4201 AAGCGGAAATGGGAGCAGCAATGATGCTGTGAAAAATATAATTTTCCCAGATGGCCC 4260
4201 AAGCGGAAATGGGAGCAGCAATGATGCTGTGAAAAATATAATTTTCCCAGATGGCCC 4260
4261 ATCAGAGCGGTTTCATTCGAACGGAAGTACAGACAGAGTTAAAGAAATAGGTTGGGAAA 4320
4261 ATCAGAGCGGTTTCATTCGAACGGAAGTACAGACAGAGTTAAAGAAATAGGTTGGGAAA 4320
4321 GAGAGCATCAAGAGAGAGAGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4380
4321 GAGAGCATCAAGAGAGAGAGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4380
4381 GCAAGTGTGGAGTATTTTACTTCTCAGTAATCTGTGACTGTGCTTATTTGAGACCTAGCCT 4440
4381 GCAAGTGTGGAGTATTTTACTTCTCAGTAATCTGTGACTGTGCTTATTTGAGACCTAGCCT 4440
4441 AGTTTCTCTGCACCAATGAACGAGTGTGCTCAATTTGAAATTAATAACAGTCAAAATCG 4500
4441 AGTTTCTCTGCACCAATGAACGAGTGTGCTCAATTTGAAATTAATAACAGTCAAAATCG 4500
4501 CTATTTGTTTAAATGATCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 4560
4501 CTATTTGTTTAAATGATCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 4560
4561 TTTCTTCTATTTAAACGGAACCTTGAATTTGGTGAATGTGATTAATTTCTTCTTATTTTCTG 4620
4561 TTTCTTCTATTTAAACGGAACCTTGAATTTGGTGAATGTGATTAATTTCTTCTTATTTTCTG 4620
4621 CTCCTTAAATTAATAAATTTCAAGAGAGCATATTTCTATGTTGGAATAGATCTGTTTTCAT 4680
4621 CTCCTTAAATTAATAAATTTCAAGAGAGCATATTTCTATGTTGGAATAGATCTGTTTTCAT 4680
4681 CTGTGCTCCAGATTTGTGACCTTAGACTTTCAATTTGACAGTGAAGTGAAGTGAAGTGAAG 4740
4681 CTGTGCTCCAGATTTGTGACCTTAGACTTTCAATTTGACAGTGAAGTGAAGTGAAGTGAAG 4740
4741 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4764
4741 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4764

RESULT 4
US-09-900-425A-1
; Sequence 1, Application US/09900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-900-425A-1

Query Match          99.9%; Score 4760.8; DB 3; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGCTTTGGTACCTGCGGTAGTACCTGCGGTGCTTCTGTCAGCGGATCTCGCGGCCGAG 60
Db 1 CTGCTTTGGTACCTGCGGTAGTACCTGCGGTGCTTCTGTCAGCGGATCTCGCGGCCGAG 60

QY 61 AGCCTTTTATAGGTGCTTTTCCCGGGGATGCTGAAGGATACAGAAATGACTGTGAATCAA 120
Db 61 AGCCTTTTATAGGTGCTTTTCCCGGGGATGCTGAAGGATACAGAAATGACTGTGAATCAA 120

QY 121 CCATATCATCAGGAGCTGATATCTAGTGGAGAGTTAGAGTGTCATCTTCACTA 180
Db 121 CCATATCATCAGGAGCTGATATCTAGTGGAGAGTTAGAGTGTCATCTTCACTA 180

QY 181 TGATATGAGGAGCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCAGGCGGA 240
Db 181 TGATATGAGGAGCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCAGGCGGA 240

QY 241 ACATATGATGCAAGGAAACACATGTACAGAAATGCTGTTCCACCGGGACGAGGCGTC 300
Db 241 ACATATGATGCAAGGAAACACATGTACAGAAATGCTGTTCCACCGGGACGAGGCGTC 300

QY 301 CCGGAGGAGGAGGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 301 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

QY 361 TGAGGCTGCTTACCTCAGCAGGCTCTGTCGAAATATCAATATGAACTTCAAGTGCCC 420
Db 361 TGAGGCTGCTTACCTCAGCAGGCTCTGTCGAAATATCAATATGAACTTCAAGTGCCC 420

QY 421 CTTCCACCACTTCTCAAACTCTCAGGCCCCCAATTTCTCCCTCCACGACGAGCTTGG 480
Db 421 CTTCCACCACTTCTCAAACTCTCAGGCCCCCAATTTCTCCCTCCACGACGAGCTTGG 480

QY 481 TACCTTTCCCCCACCATGCTCCTGTCAGCGCAAGGCGCTCTTCCCTCCGAGGAGGAGG 540
Db 481 TACCTTTCCCCCACCATGCTCCTGTCAGCGCAAGGCGCTCTTCCCTCCGAGGAGGAGG 540

QY 541 GCGCGCTTTCCGCAACCAACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 541 GCGCGCTTTCCGCAACCAACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600

QY 601 CCATGCGCACCAACCAATGCTTGTCTTAATAACCCCGAGTCCCTGGGCGACCTCTGGAC 660
Db 601 CCATGCGCACCAACCAATGCTTGTCTTAATAACCCCGAGTCCCTGGGCGACCTCTGGAC 660

QY 661 AAGGCACTTTCCCTTCATGATGCGCCCTCCTCCATGCTCATCCCGCGCCCTCCAG 720
Db 661 AAGGCACTTTCCCTTCATGATGCGCCCTCCTCCATGCTCATCCCGCGCCCTCCAG 720

QY 721 TCATGCGGAGGAGGTTAATATCATGACCTTCCGCGGCTTCTCAACCAACTTCCAC 780
Db 721 TCATGCGGAGGAGGTTAATATCATGACCTTCCGCGGCTTCTCAACCAACTTCCAC 780

QY 781 CTCGCCAGTTTAAATAGTTTCCAGAACACCTTAGTTCTTCTGCGCCAGTGTATATCA 840
Db 781 CTCGCCAGTTTAAATAGTTTCCAGAACACCTTAGTTCTTCTGCGCCAGTGTATATCA 840

QY 841 CGAGTAGTCTCATTTACAGATCTCTCTCCATACCACTCCCTCCAAAGGCTCCAGTGAGA 900
Db 841 CGAGTAGTCTCATTTACAGATCTCTCTCCATACCACTCCCTCCAAAGGCTCCAGTGAGA 900

QY 901 GAAGGTCCCCAGAAAGGCTGAAACACTATGATGACCAAGGAGGAGGAGGAGGAGGAGG 960
Db 901 GAAGGTCCCCAGAAAGGCTGAAACACTATGATGACCAAGGAGGAGGAGGAGGAGGAGG 960
```

[illegible]

[illegible]

RESULT 5

```

US-10-805-919-1
; Sequence 1, Application US/10805919
; Publication No. US20040175828A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/10/805,919
; CURRENT FILING DATE: 2004-03-22
; PRIOR APPLICATION NUMBER: US/09/900,425
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-805-919-1

Query Match          99.9%; Score 4760.8; DB 8; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY      1      CTGCTCTGGTACTCTGCGGTAGTACCTTGGCTTTTGCTCTGACGGCGATCTCGCGGCCGAG 60

```


QY	1141	GATCACAATCTCTGGAAAGGTCCTCAAAAAGAGTATAGAGATCTGGAAGGATTACG	1200
DB	1141	GATCACAATCTCTGGAAAGGTCCTCAAAAAGAGTATAGAGATCTGGAAGGATTACG	1200
QY	1201	GTTTATCGGTGTTCTCGAACCTCTGGATGCACACAGAAATACCTGGGGAGATTATTA	1260
DB	1201	GTTTATCGGTGTTCTCGAACCTCTGGATGCACACAGAAATACCTGGGGAGATTATTA	1260
QY	1261	AAAATACAGATTCCTGGGCCCCCCTCGAGATTGTGAATCATCGTCCCCAAGTAGGG	1320
DB	1261	AAAATACAGATTCCTGGGCCCCCCTCGAGATTGTGAATCATCGTCCCCAAGTAGGG	1320
QY	1321	AGAAAGAGAGCTCGTTGGGAGGAGAAAAGACCGTGGAGTGACACACAGAGTTCTG	1380
DB	1321	AGAAAGAGAGCTCGTTGGGAGGAGAAAAGACCGTGGAGTGACACACAGAGTTCTG	1380
QY	1381	GCAAGACAGAACTATACCTCAATCAAGGAAAAGAGCCGAGGAGACCATCCCTGACA	1440
DB	1381	GCAAGACAGAACTATACCTCAATCAAGGAAAAGAGCCGAGGAGACCATCCCTGACA	1440
QY	1441	AGATGAGGAGGAAGAAAGAACTTCTTAAGCCTGTGTGGATTTCGATGCACCTCATTCAG	1500
DB	1441	AGATGAGGAGGAAGAAAGAACTTCTTAAGCCTGTGTGGATTTCGATGCACCTCATTCAG	1500
QY	1501	AAAACTACTACTCCAGTGCACCAATGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA	1560
DB	1501	AAAACTACTACTCCAGTGCACCAATGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA	1560
QY	1561	GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG	1620
DB	1561	GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG	1620
QY	1621	CCAAAGCTGCTCGGCCCTCGTGGAACTCCAAAGACGAAGCTCGATGAAGATTAGAGA	1680
DB	1621	CCAAAGCTGCTCGGCCCTCGTGGAACTCCAAAGACGAAGCTCGATGAAGATTAGAGA	1680
QY	1681	GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACGACCACTGTTCTAGCAGCTCAGACT	1740
DB	1681	GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACGACCACTGTTCTAGCAGCTCAGACT	1740
QY	1741	CTGAAGTTTTGAGTTTTGCGAAGTCAAAAGCAAAAGGCCCCCTGACCGACTTC	1800
DB	1741	CTGAAGTTTTGAGTTTTGCGAAGTCAAAAGCAAAAGGCCCCCTGACCGACTTC	1800
QY	1801	ATGATGAACCTTTGGTCAACAGATCCAGGCCAGATGAATGATGACCACTCTGCAAAATGCA	1860
DB	1801	ATGATGAACCTTTGGTCAACAGATCCAGGCCAGATGAATGATGACCACTCTGCAAAATGCA	1860
QY	1861	GCGCAAGGCAAGACGACAGGAATTAGGCAAGCATTTATCTGGAGAGAGGCCATCA	1920
DB	1861	GCGCAAGGCAAGACGACAGGAATTAGGCAAGCATTTATCTGGAGAGAGGCCATCA	1920
QY	1921	AGCCCTGTCTCTATGACCAACATGCTGGCAGACTTTTCCATACCGGATCACAGTCT	1980
DB	1921	AGCCCTGTCTCTATGACCAACATGCTGGCAGACTTTTCCATACCGGATCACAGTCT	1980
QY	1981	CCCGCCTACCAACTTTTAACTGACAGGCCAATGTTATAGATACGATGATCACCGAGT	2040
DB	1981	CCCGCCTACCAACTTTTAACTGACAGGCCAATGTTATAGATACGATGATCACCGAGT	2040
QY	2041	ATATCTTTGAAGGATTTTCTATGTTTGCATGCCCCCTGACCAATATTCACCTGTGTA	2100
DB	2041	ATATCTTTGAAGGATTTTCTATGTTTGCATGCCCCCTGACCAATATTCACCTGTGTA	2100
QY	2101	AAGTAATTAGATTCAACATAGACTACAGGATTCATTTCAATGGAAGATGATGCCGAGA	2160
DB	2101	AAGTAATTAGATTCAACATAGACTACAGGATTCATTTCAATGGAAGATGATGCCGAGA	2160
QY	2161	ATTTTGTGTCAAAAGGGCTGAACTCTTTTCACTGTTCTCTATTACAGAGATATTTTGAAT	2220
DB	2161	ATTTTGTGTCAAAAGGGCTGAACTCTTTTCACTGTTCTCTATTACAGAGATATTTTGAAT	2220

QY	2221	TATATGATCGGAATCTTAAAGGTCCTTTGTTTGAAGACAGCCCTCCCTGCTGCCCCAAGAT	2280
DB	2221	TATATGATCGGAATCTTAAAGGTCCTTTGTTTGAAGACAGCCCTCCCTGCTGCCCCAAGAT	2280
QY	2281	TTCAATTTCAATGCGCAAGTTTGTAAAGTTCTTCCAGATGGAGAAAGGAAGTCTGTCCA	2340
DB	2281	TTCAATTTCAATGCGCAAGTTTGTAAAGTTCTTCCAGATGGAGAAAGGAAGTCTGTCCA	2340
QY	2341	TGCACCAAGATTTCTCTGTACTTTTAAAGTGCAGCAAAAGCCCTGGTGCCTGAGGAGAGA	2400
DB	2341	TGCACCAAGATTTCTCTGTACTTTTAAAGTGCAGCAAAAGCCCTGGTGCCTGAGGAGAGA	2400
QY	2401	TTGCCAATATGCTTCAAGTGGGAGGAGCTGGAGTGGCAGAAAATATGAGAGAAATGCAAAAG	2460
DB	2401	TTGCCAATATGCTTCAAGTGGGAGGAGCTGGAGTGGCAGAAAATATGAGAGAAATGCAAAAG	2460
QY	2461	GCATGATTTGTTACCAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC	2520
DB	2461	GCATGATTTGTTACCAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC	2520
QY	2521	GTGACAGTTTCAACCCCGATGTGATTACTTTTCCGATTATCGTCCACTTTTGGGATACGCC	2580
DB	2521	GTGACAGTTTCAACCCCGATGTGATTACTTTTCCGATTATCGTCCACTTTTGGGATACGCC	2580
QY	2581	CTGCACAGTTGAGTTATGACAGGAGACCCACAGTACCAAAAATCTGTGGAAAGATTATGTGA	2640
DB	2581	CTGCACAGTTGAGTTATGACAGGAGACCCACAGTACCAAAAATCTGTGGAAAGATTATGTGA	2640
QY	2641	AACCTTCCCACTCTTAGCAAAATAGTCCCAAAAGTCAAAACAACTGACAAACAGAAAGCTGG	2700
DB	2641	AACCTTCCCACTCTTAGCAAAATAGTCCCAAAAGTCAAAACAACTGACAAACAGAAAGCTGG	2700
QY	2701	CACAGAGGAGGAGGAGCCCTCCAAAATATGCGCAGAGAAATACAATGAGACGAGAAAGTAA	2760
DB	2701	CACAGAGGAGGAGGAGCCCTCCAAAATATGCGCAGAGAAATACAATGAGACGAGAAAGTAA	2760
QY	2761	CGGTGAGCTTAAGTACGCAAGGATTTGGAATACTGGCAATCTGTCATCTGTGATGTCTGCAGC	2820
DB	2761	CGGTGAGCTTAAGTACGCAAGGATTTGGAATACTGGCAATCTGTCATCTGTGATGTCTGCAGC	2820
QY	2821	ATGCAATGATGTCTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCTTAATGCAAT	2880
DB	2821	ATGCAATGATGTCTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCTTAATGCAAT	2880
QY	2881	TGACCAAGTTGATAGATATATCTTCCAGATCGTGTCTGTTGCAAGCTGCGCATGACTC	2940
DB	2881	TGACCAAGTTGATAGATATATCTTCCAGATCGTGTCTGTTGCAAGCTGCGCATGACTC	2940
QY	2941	ATCCAGTCAATCAATTTAAATTTTGGAAATGAATCTGATCATGCGAGGAATTCATTAATCTA	3000
DB	2941	ATCCAGTCAATCAATTTAAATTTTGGAAATGAATCTGATCATGCGAGGAATTCATTAATCTA	3000
QY	3001	ACTGTGGAATTTCCGACGCCCAATATGAGGAGACAGAAAGTTTCATCATGACATGCGGA	3060
DB	3001	ACTGTGGAATTTCCGACGCCCAATATGAGGAGACAGAAAGTTTCATCATGACATGCGGA	3060
QY	3061	AGAAAGGATTAACAACCTTGCATATAATCATGTGACGCCCTTGGCAGAGTACCCCACTC	3120
DB	3061	AGAAAGGATTAACAACCTTGCATATAATCATGTGACGCCCTTGGCAGAGTACCCCACTC	3120
QY	3121	CCTCGAGGATTAACAACCTTGGAAATTCCTGGGTGATGCTGTTGTTGAAATTC	3180
DB	3121	CCTCGAGGATTAACAACCTTGGAAATTCCTGGGTGATGCTGTTGTTGAAATTC	3180
QY	3181	TGACACCGCTCCATTTGTACTATTTGTTTCTAGTCTGGAAGAGAGGAGATTAGCAACCT	3240
DB	3181	TGACACCGCTCCATTTGTACTATTTGTTTCTAGTCTGGAAGAGAGGAGATTAGCAACCT	3240
QY	3241	ATCGGATGCGATTTGTTAGAAATCAGCACTTGCCTAGTACCAAGAACTTCAACTGG	3300
DB	3241	ATCGGATGCGATTTGTTAGAAATCAGCACTTGCCTAGTACCAAGAACTTCAACTGG	3300
QY	3301	ATCCATTTATGCTGATGCTCAACGCGGCTGACCTTTTGTAGAGATCGGACCTTCGACATG	3360

[illegible]

Db	4361	GCAAGTGTGGAGTAATTTA	CTTGCTCAGTAAC	TGTCGACTGTTGCTATT	TGAGACCTAGCC	4444
Qy	4441	AGTTTTCTCGACAACAATGA	CAGAAGTGTGCTCAAT	GAAATAAAATACAGAGTCAAATCG	4500	
Db	4441	AGTTTTCTCGACAACAATGA	CAGAAGTGTGCTCAAT	GAAATAAAATACAGAGTCAAATCG	4500	
Qy	4501	CTATTGTTGTTTTAAATGAT	CTGTTTTTAGCTGGAGTGC	TTTATTA	CAAAAGTATTAGATT	4560
Db	4501	CTATTGTTGTTTTAAATGAT	CTGTTTTTAGCTGGAGTGC	TTTATTA	CAAAAGTATTAGATT	4560
Qy	4561	TTTCTTCTATTTAACGGAA	AACTTGACTTTGGTGAATG	TGCATTACTTCTCTTTATTTTG	4620	
Db	4561	TTTCTTCTATTTAACGGAA	AACTTGACTTTGGTGAATG	TGCATTACTTCTCTTTATTTTG	4620	
Qy	4621	CTCTTTAAATAATAAAAA	TTCAAGAAGCATATTTCTAT	GTGGAATAGATCCTGTTTTTCCAT	4680	
Db	4621	CTCTTTAAATAATAAAAA	TTCAAGAAGCATATTTCTAT	GTGGAATAGATCCTGTTTTTCCAT	4680	
Qy	4681	CTGTGTCCAGANTGTGACC	CTAGACTTTCAATTGACAAG	TAAATAATGACCTTTTACTAG	4740	
Db	4681	CTGTGTCCAGANTGTGACC	CTAGACTTTCAATTGACAAG	TAAATAATGACCTTTTACTAG	4740	
Qy	4741	TAAAAA	AAAAAAAAAAAAAAAAAAAA	A	4764	
Db	4741	TAAAAA	AAAAAAAAAAAAAAAAAAAA	A	4764	
RESULT 6						
US-10-723-860-280						
; Sequence 280, Application US/10723860						
; Publication No. US20040253606A1						
; GENERAL INFORMATION:						
; APPLICANT: Aziz, Natasha						
; APPLICANT: Ginsburg, Wendy M.						
; APPLICANT: Zlotnik, Albert						
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Composition						
; FILE REFERENCE: 05882.0193.NPUS01						
; CURRENT APPLICATION NUMBER: US/10723,860						
; CURRENT FILING DATE: 2003-11-26						
; PRIOR APPLICATION NUMBER: 60/429,739						
; PRIOR FILING DATE: 2002-11-26						
; NUMBER OF SEQ ID NOS: 8393						
; SOFTWARE: PatentIn version 3.2						
; SEQ ID NO 280						
; LENGTH: 4764						
; TYPE: DNA						
; ORGANISM: Homo sapiens						
US-10-723-860-280						
Query Match 99.9%; Score 4760.8; DB 8; Length 4764;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;						
Qy	1	CTGTCTTGGTACCTGCGGTAGTAC	CGCTTGGCTTTGCTCTGACGCGCATCT	CGCGGCCGAG	60	
Db	1	CTGTCTTGGTACCTGCGGTAGTAC	CGCTTGGCTTTGCTCTGACGCGCATCT	CGCGGCCGAG	60	
Qy	61	AGCCTTTTATAGTTTGTCTTTT	CCGGGGATGTGAAGGATACAGAA	TGACTGTGGAATCAA	120	
Db	61	AGCCTTTTATAGTTTGTCTTTT	CCGGGGATGTGAAGGATACAGAA	TGACTGTGGAATCAA	120	
Qy	121	CCCATATCATCAAGGAGCTGAT	AATCTAGTGGGAAGATTAGAC	GTCTGCATATCTTCCACTA	180	
Db	121	CCCATATCATCAAGGAGCTGAT	AATCTAGTGGGAAGATTAGAC	GTCTGCATATCTTCCACTA	180	
Qy	181	TGATATGAGGCAGTCTCTGAG	CTTATATTTCTCTGTGGAAAGAT	GTGTGACATATCCAGCGGA	240	
Db	181	TGATATGAGGCAGTCTCTGAG	CTTATATTTCTCTGTGGAAAGAT	GTGTGACATATCCAGCGGA	240	
Qy	241	ACATCATGATGCAGGAAAACA	ATGTACAGAAATGTGTTTCCA	CCCGGACGAGGGCGTTC	300	
Db	241	ACATCATGATGCAGGAAAACA	ATGTACAGAAATGTGTTTCCA	CCCGGACGAGGGCGTTC	300	

QY 301 CCCGAGCAGGAGGACATGGAGCCAGACCCCTCAGACAATCTTTTAGGCCCCAAAATC 360
Db |||||
QY 301 CCCGAGCAGGAGGACATGGAGCCAGACCCCTCAGACAATCTTTTAGGCCCCAAAATC 360
Db |||||
QY 361 TGAAGGTGCTTACCCCTCAGCAGCCTCTGTGCAATATCAATATGAACCTTCAAGTGCC 420
Db |||||
QY 361 TGAAGGTGCTTACCCCTCAGCAGCCTCTGTGCAATATCAATATGAACCTTCAAGTGCC 420
Db |||||
QY 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCAGCAGCAGACTTTG 480
Db |||||
QY 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCAGCAGCAGACTTTG 480
Db |||||
QY 481 TACCTTTCCCCCAACCCATGCTCCGTGAGCGAAGGCCCTCTTCCCTCCCTGCCCATA 540
Db |||||
QY 481 TACCTTTCCCCCAACCCATGCTCCGTGAGCGAAGGCCCTCTTCCCTCCCTGCCCATA 540
Db |||||
QY 541 GGCCTTTCCCAACCCATGAGGACCCCTTCCAGTTCCTCTGTTGTTTCCCTC 600
Db |||||
QY 541 GGCCTTTCCCAACCCATGAGGACCCCTTCCAGTTCCTCTGTTGTTTCCCTC 600
Db |||||
QY 601 CCATGCCACCACTTCTGCTTAAATACCCCTCCAGTCCCTGGGGACCTCTTGAC 660
Db |||||
QY 601 CCATGCCACCACTTCTGCTTAAATACCCCTCCAGTCCCTGGGGACCTCTTGAC 660
Db |||||
QY 661 AAGCCTTTCCCTTCCATGATGCCCCCTCCCTCCATGCTCCCTCCCTCCCTCCAG 720
Db |||||
QY 661 AAGCCTTTCCCTTCCATGATGCCCCCTCCCTCCATGCTCCCTCCCTCCCTCCAG 720
Db |||||
QY 721 TCATGCCGAGCGAGTTAATATCAGTACCCCTGGGCTATTCTCACCACAACTTCCCAC 780
Db |||||
QY 721 TCATGCCGAGCGAGTTAATATCAGTACCCCTGGGCTATTCTCACCACAACTTCCCAC 780
Db |||||
QY 781 CTCCAGTTTAAATAGTTTCCAGAACCACTAGTTCTTCTTCCCTGCCCAGTCTAAATA 840
Db |||||
QY 781 CTCCAGTTTAAATAGTTTCCAGAACCACTAGTTCTTCTTCCCTGCCCAGTCTAAATA 840
Db |||||
QY 841 GCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCAGTGAGA 900
Db |||||
QY 841 GCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCAGTGAGA 900
Db |||||
QY 901 GAAGGTCCCAGAAAGGCTGAAACACTATGATGACCAAGCCAGCAGCAGCAGTCTATG 960
Db |||||
QY 901 GAAGGTCCCAGAAAGGCTGAAACACTATGATGACCAAGCCAGCAGCAGCAGTCTATG 960
Db |||||
QY 961 GGCAGGTGAGGAGCATCGGTCCCTGGATCGCGGAGCGAGGCGCAGTCCCGACAGGA 1020
Db |||||
QY 961 GGCAGGTGAGGAGCATCGGTCCCTGGATCGCGGAGCGAGGCGCAGTCCCGACAGGA 1020
Db |||||
QY 1021 GAAGCAAGCAGCGGTACAGATCTGATTATGACCGAGGAGAACCAATCTCGCCACC 1080
Db |||||
QY 1021 GAAGCAAGCAGCGGTACAGATCTGATTATGACCGAGGAGAACCAATCTCGCCACC 1080
Db |||||
QY 1081 GCAGCTACGAACGAGCAGAGCGAGAAACCGGAGAGACACAGGATCGAGACACCGAA 1140
Db |||||
QY 1081 GCAGCTACGAACGAGCAGAGCGAGAAACCGGAGAGACACAGGATCGAGACACCGAA 1140
Db |||||
QY 1141 GATCACCATCTCTGGAAGGTCCTACAAAAGAGTATAAGAGATCTGGAAGGAGTTACG 1200
Db |||||
QY 1141 GATCACCATCTCTGGAAGGTCCTACAAAAGAGTATAAGAGATCTGGAAGGAGTTACG 1200
Db |||||
QY 1201 GTTTATCGGTTGTTCTGAACTCTGATGACACCAAGATTTACTCTGGGAGATTATTA 1260
Db |||||
QY 1201 GTTTATCGGTTGTTCTGAACTCTGATGACACCAAGATTTACTCTGGGAGATTATTA 1260
Db |||||
QY 1261 AAAATACAGATTTCTTGGGCCCAACCCCTGGAGATTGTGAATCATGCTCCCAAGTAGGG 1320
Db |||||
QY 1261 AAAATACAGATTTCTTGGGCCCAACCCCTGGAGATTGTGAATCATGCTCCCAAGTAGGG 1320
Db |||||
QY 1321 AGAAGAGAGCTGTTGGGAGGAGAAAGACCGTTGGAGTGACAAACAGAGTTCTG 1380
Db |||||
QY 1321 AGAAGAGAGAGCTGTTGGGAGGAGAAAGACCGTTGGAGTGACAAACAGAGTTCTG 1380
Db |||||

QY 1381 GCAAAGACAAGAACTATACCTCAATCAAGGAAAAAGAGCCCGAGGAGACCATGCTCTGACA 1440
Db |||||
QY 1381 GCAAAGACAAGAACTATACCTCAATCAAGGAAAAAGAGCCCGAGGAGACCATGCTCTGACA 1440
Db |||||
QY 1441 AGAATGAGGAGGAAGAAAGAACTTTTAAAGCTGTGTGGATTTCATGATCCTCATTTAG 1500
Db |||||
QY 1441 AGAATGAGGAGGAAGAAAGAACTTTTAAAGCTGTGTGGATTTCATGATCCTCATTTAG 1500
Db |||||
QY 1501 AAAATCTACTCTCCAGTGACCCCATGGATCAGTGGGAGATTCTACAGTGGTTGGAACGA 1560
Db |||||
QY 1501 AAAATCTACTCTCCAGTGACCCCATGGATCAGTGGGAGATTCTACAGTGGTTGGAACGA 1560
Db |||||
QY 1561 GTAGGCTTCTGCTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG 1620
Db |||||
QY 1561 GTAGGCTTCTGCTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG 1620
Db |||||
QY 1621 CCAAAGCTCTCGGCCCTCGTGGGAACCTCCAAAAGACGAAGCTCGATGAGATTATAGAGA 1680
Db |||||
QY 1621 CCAAAGCTCTCGGCCCTCGTGGGAACCTCCAAAAGACGAAGCTCGATGAGATTATAGAGA 1680
Db |||||
QY 1681 GTTCCAGTGAATCCCGAGTGTGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1740
Db |||||
QY 1681 GTTCCAGTGAATCCCGAGTGTGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1740
Db |||||
QY 1741 CTGAAGTTTTGACGTTATTGACAGAAATCAAAACGCAAAAGGCCCAACCTGACCGACTTC 1800
Db |||||
QY 1741 CTGAAGTTTTGACGTTATTGACAGAAATCAAAACGCAAAAGGCCCAACCTGACCGACTTC 1800
Db |||||
QY 1801 ATGATGAACCTTTGGTTCAACAGATCCAGGCCAGATGAATGATGAGCACTCTGCAAAATGCA 1860
Db |||||
QY 1801 ATGATGAACCTTTGGTTCAACAGATCCAGGCCAGATGAATGATGAGCACTCTGCAAAATGCA 1860
Db |||||
QY 1861 GCGCAAGGACAGACCGCAGGAAATTTAGGACAGCACTTTATCTCGGAGAGAGGCCATCA 1920
Db |||||
QY 1861 GCGCAAGGACAGACCGCAGGAAATTTAGGACAGCACTTTATCTCGGAGAGAGGCCATCA 1920
Db |||||
QY 1921 AGCCCTGTCTGCTATGACCAACAAATGCTGGGAGAGATTTTCCACTACCGGATCACAGTCT 1980
Db |||||
QY 1921 AGCCCTGTCTGCTATGACCAACAAATGCTGGGAGAGATTTTCCACTACCGGATCACAGTCT 1980
Db |||||
QY 1981 CCGCGCTACGAACCTTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCACAGT 2040
Db |||||
QY 1981 CCGCGCTACGAACCTTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCACAGT 2040
Db |||||
QY 2041 ATATCTTTGAAAGATTTTCTATGTTGCACTGCCCCCTCGACCAATTTCCACTGTGTA 2100
Db |||||
QY 2041 ATATCTTTGAAAGATTTTCTATGTTGCACTGCCCCCTCGACCAATTTCCACTGTGTA 2100
Db |||||
QY 2101 AAGTAAATTTAGATTTCAACATAGACTACACGATTCATTTCAATTGAAAGAGATGATGCCGAGA 2160
Db |||||
QY 2101 AAGTAAATTTAGATTTCAACATAGACTACACGATTCATTTCAATTGAAAGAGATGATGCCGAGA 2160
Db |||||
QY 2161 ATTTTGTGTGAAAGGCTTTGAACTCTTTTCACTGTTCCTATTTCAGAGATATTTTGAAT 2220
Db |||||
QY 2161 ATTTTGTGTGAAAGGCTTTGAACTCTTTTCACTGTTCCTATTTCAGAGATATTTTGAAT 2220
Db |||||
QY 2221 TATATGACTGGAATCTTAAAGGCTTTTGTGAGACAGCCCTCCCTGCTGCCCAAGAT 2280
Db |||||
QY 2221 TATATGACTGGAATCTTAAAGGCTTTTGTGAGACAGCCCTCCCTGCTGCCCAAGAT 2280
Db |||||
QY 2281 TTTCAATTTTCAGCCACGTTTTTGTAAAGATTTCTTCAGATGGAGAAAGAGATGCTGTCCA 2340
Db |||||
QY 2281 TTTCAATTTTCAGCCACGTTTTTGTAAAGATTTCTTCAGATGGAGAAAGAGATGCTGTCCA 2340
Db |||||
QY 2341 TGCAACAGATTTCTCTGTACTTTGTTAAAGGTGACGCAAGGCCCTGGTCTGAGAGGAGA 2400
Db |||||
QY 2341 TGCAACAGATTTCTCTGTACTTTGTTAAAGGTGACGCAAGGCCCTGGTCTGAGAGGAGA 2400
Db |||||
QY 2401 TTGCAATATGCTTTTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAG 2460
Db |||||
QY 2401 TTGCAATATGCTTTTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAG 2460
Db |||||
QY 2461 GCATGATTTGTACCAACCCCTGGGAGGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
Db |||||

Db	2461	GCATGATTTGTTTACAAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGCGATC	2520
Qy	2521	GTGAACAGTTCAAACCCGATGTGATTACTTTTCCGATTTATCGTCCACTTTTGGGATAGCC	2580
Db	2521	GTGAACAGTTCAAACCCGATGTGATTACTTTTCCGATTTATCGTCCACTTTTGGGATAGCC	2580
Qy	2581	CTGCACAGTTGAGTTATGCAAGGAGACCCACAGTACCAAAAACCTGTGGAAGAGTTATGTGA	2640
Db	2581	CTGCACAGTTGAGTTATGCAAGGAGACCCACAGTACCAAAAACCTGTGGAAGAGTTATGTGA	2640
Qy	2641	AACTTCCGCCACTCCTTAGCAATAAGTCCCAAGTCAAACTGACAAACAGNAGCTGG	2700
Db	2641	AACTTCCGCCACTCCTTAGCAATAAGTCCCAAGTCAAACTGACAAACAGNAGCTGG	2700
Qy	2701	CACAGAGGAGGAAGCCCTCCAAAAATACGGCAGAAGNATACAATCAGACACGAGAAGTAA	2760
Db	2701	CACAGAGGAGGAAGCCCTCCAAAAATACGGCAGAAGNATACAATCAGACACGAGAAGTAA	2760
Qy	2761	CGGTGGAGCTAAGTAGCAAGGATTTCTGAAAACTGCGCATCCGTTCTGTCTGTCAGC	2820
Db	2761	CGGTGGAGCTAAGTAGCAAGGATTTCTGAAAACTGCGCATCCGTTCTGTCTGTCAGC	2820
Qy	2821	ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAATGCTTAATGCAATT	2880
Db	2821	ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAATGCTTAATGCAATT	2880
Qy	2881	TGGACAAGTTGATAGGATATACTTTCCAAGATCGTGTCTGTTGTGACGTGCGCATGACTC	2940
Db	2881	TGGACAAGTTGATAGGATATACTTTCCAAGATCGTGTCTGTTGTGACGTGCGCATGACTC	2940
Qy	2941	ATCCAAGTCACTAATTTAAAATTTTGGAAATGAATCCTGTATCATGCGCAGGAATTCATTATCTA	3000
Db	2941	ATCCAAGTCACTAATTTAAAATTTTGGAAATGAATCCTGTATCATGCGCAGGAATTCATTATCTA	3000
Qy	3001	ACTGTGGAATTCGGCAGCCCAATAACGGAGACAGAAAAGTTTCATCATGCAATGCGGA	3060
Db	3001	ACTGTGGAATTCGGCAGCCCAATAACGGAGACAGAAAAGTTTCATCATGCAATGCGGA	3060
Qy	3061	AGAAAGGATTAACACCTTCGATTAATATCATGTCAAGCTTTGGCAGAGATGACCCAACTC	3120
Db	3061	AGAAAGGATTAACACCTTCGATTAATATCATGTCAAGCTTTGGCAGAGATGACCCAACTC	3120
Qy	3121	CCTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTGTTGAATTTTC	3180
Db	3121	CCTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTGTTGAATTTTC	3180
Qy	3181	TGACCAAGCTCCATTTGTAATAATTTGTTTCTTAGTCTGGAAGAGGAGGATTAGCAACCT	3240
Db	3181	TGACCAAGCTCCATTTGTAATAATTTGTTTCTTAGTCTGGAAGAGGAGGATTAGCAACCT	3240
Qy	3241	ATCGGACTGCCATTTGTCAGAAATCAGCACCTGCCATGCTAGCAAGAACTTGAACCTG	3300
Db	3241	ATCGGACTGCCATTTGTCAGAAATCAGCACCTGCCATGCTAGCAAGAACTTGAACCTG	3300
Qy	3301	ATCCATTTATGCTGTATGCTCAGGGGCTGACACTTTGTAGAGAATCGGACCTTCGACATG	3360
Db	3301	ATCCATTTATGCTGTATGCTCAGGGGCTGACACTTTGTAGAGAATCGGACCTTCGACATG	3360
Qy	3361	CAATGCGCAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGGAAGCCTGGAGG	3420
Db	3361	CAATGCGCAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGGAAGCCTGGAGG	3420
Qy	3421	AAGCCAGCAGTTATTTGGACGCTTGCTCTTTAATGATCCGGACCTCGCGAAGTCTGGC	3480
Db	3421	AAGCCAGCAGTTATTTGGACGCTTGCTCTTTAATGATCCGGACCTCGCGAAGTCTGGC	3480
Qy	3481	TCAATTAATCCTCTCCACCACTCCAATCAAGAGCCAAATATCTGATCGACAACCTTATTG	3540
Db	3481	TCAATTAATCCTCTCCACCACTCCAATCAAGAGCCAAATATCTGATCGACAACCTTATTG	3540
Qy	3541	AAACTTCTCCAGTTCTCAAAAACTTACTGAGTTTGAAGAAGCAATTTGGAGTAAATTTTAA	3600

QY 4681 CTGTGTCAGATGTGACCCCTAGACTTTCAATTGACAAAGTAAAAAATTGACTTTACTAG 4740
DB |||||||
QY 4681 CTGTGTCAGATGTGACCCCTAGACTTTCAATTGACAAAGTAAAAAATTGACTTTACTAG 4740
DB |||||||
QY 4741 TAAAAAATAAAAAAAAAAAAAA 4764
DB |||||||
QY 4741 TAAAAAATAAAAAAAAAAAAAA 4764
DB |||||||

RESULT 7
US-10-756-149-319
; Sequence 319, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 319
; LENGTH: 4764
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-319

Query Match 99.9%; Score 4760.8; DB 9; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CTGTCTTGTTGACCTGCGGTAGTACCTGGCTTTGCTCTGACGGCGATCTCGCGGCCGAG 60
DB 1 CTGTCTTGTTGACCTGCGGTAGTACCTGGCTTTGCTCTGACGGCGATCTCGCGGCCGAG 60

QY 61 AGCCCTTTATAGTGTCTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
DB 61 AGCCCTTTATAGTGTCTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
QY 121 CCCATATCATCAAGGAGCTGATAATCTAGTGGAAAGATTAGACGTGTGCATATCTTCACTA 180
DB 121 CCCATATCATCAAGGAGCTGATAATCTAGTGGAAAGATTAGACGTGTGCATATCTTCACTA 180
QY 181 TGATATGAGGAGCTCTGAGCTTATATTCTCTGTGGAAGATGTGACATATCCAGGCGGA 240
DB 181 TGATATGAGGAGCTCTGAGCTTATATTCTCTGTGGAAGATGTGACATATCCAGGCGGA 240
QY 241 ACATCATGATCGAGGAAACACATGTCAAGAAATGTCTTCCACCCGCGGAGCGGGGCTC 300
DB 241 ACATCATGATCGAGGAAACACATGTCAAGAAATGTCTTCCACCCGCGGAGCGGGGCTC 300
QY 301 CCCGAGGACGAGGAGGACATGAGCCAGACCTCTGACCATCTCTTTAGGCCCCCAAAATC 360
DB 301 CCCGAGGACGAGGAGGACATGAGCCAGACCTCTGACCATCTCTTTAGGCCCCCAAAATC 360
QY 361 TGAGGCTGCTTCACTCCTCAGAGCTCTCTGTGCAATATCAATATGAACTTCAAGTGCC 420
DB 361 TGAGGCTGCTTCACTCCTCAGAGCTCTCTGTGCAATATCAATATGAACTTCAAGTGCC 420
QY 421 CTTCCACCACTTTCTCAAACTCTCAGGCCCCCAATTTTCTCCCTCCAGCACGACATTTG 480
DB 421 CTTCCACCACTTTCTCAAACTCTCAGGCCCCCAATTTTCTCCCTCCAGCACGACATTTG 480
QY 481 TACCTCTCCGCCACCCATGCTCTGTCAGGCAAGGCGCTCTTCCGCCCTGCCCAATCA 540
DB 481 TACCTCTCCGCCACCCATGCTCTGTCAGGCAAGGCGCTCTTCCGCCCTGCCCAATCA 540
QY 541 GCGCGCTTTTCCCAACCAACGAGATGAGGACCCCTTCCAGTTCCTCTGTTTCTC 600
DB 541 GCGCGCTTTTCCCAACCAACGAGATGAGGACCCCTTCCAGTTCCTCTGTTTCTC 600

QY 601 CCATGCCACCAATGCTTGTCTTAATAAATCCCTCAGTCCCTGGGGCACCTCTCTGAC 660
DB |||||||
QY 601 CCATGCCACCAATGCTTGTCTTAATAAATCCCTCAGTCCCTGGGGCACCTCTCTGAC 660
DB |||||||
QY 661 AAGGCACTTTTCCCTTTCATGATGCCCCCTCTCATGCTCATCCCGGCCCTCCAG 720
DB |||||||
QY 661 AAGGCACTTTTCCCTTTCATGATGCCCCCTCTCATGCTCATCCCGGCCCTCCAG 720
DB |||||||
QY 721 TCATGCGGAGCAGGTTTAATTATCAGTACCTTCCGGGCTATTCTCACCACAATTTCCAC 780
DB |||||||
QY 721 TCATGCGGAGCAGGTTTAATTATCAGTACCTTCCGGGCTATTCTCACCACAATTTCCAC 780
DB |||||||
QY 781 CTCCAGTTTTAATAGTTTCCAGAAACAACCTTAGTTCTTCTGCCCCAGTGTATAACA 840
DB |||||||
QY 781 CTCCAGTTTTAATAGTTTCCAGAAACAACCTTAGTTCTTCTGCCCCAGTGTATAACA 840
DB |||||||
QY 841 GCAGTAGTCTCTCATTTTCAGACATCTCCCTCCATACCCATCTCCCAAGGCTCCAGTGAG 900
DB |||||||
QY 841 GCAGTAGTCTCTCATTTTCAGACATCTCCCTCCATACCCATCTCCCAAGGCTCCAGTGAG 900
DB |||||||
QY 901 GAAGTCCCCAGAAAGGCTGAAACACATATGATGACACAGGACCGAGACCAAGTCAATG 960
DB |||||||
QY 901 GAAGTCCCCAGAAAGGCTGAAACACATATGATGACACAGGACCGAGACCAAGTCAATG 960
DB |||||||
QY 961 GCGGAGTGCAGAGGCTCGGTCCCTGGATCGCGGGAGCGAGCGCGAGTCCCGACAGGA 1020
DB |||||||
QY 961 GCGGAGTGCAGAGGCTCGGTCCCTGGATCGCGGGAGCGAGCGCGAGTCCCGACAGGA 1020
DB |||||||
QY 1021 GAAGCAAGACAGCCGGTACAGATCTGATATGACCGGGAGAACACCATCTCGCCACC 1080
DB |||||||
QY 1021 GAAGCAAGACAGCCGGTACAGATCTGATATGACCGGGAGAACACCATCTCGCCACC 1080
DB |||||||
QY 1081 GCAGTACGAAACGAGACGAGACGAGACGAGGACACAGGATCTCAGACAAACCGAA 1140
DB |||||||
QY 1081 GCAGTACGAAACGAGACGAGACGAGACGAGGACACAGGATCTCAGACAAACCGAA 1140
DB |||||||
QY 1141 GATCCATCTCTGGAAGGCTCTCAAAAAAGATTAAGAGATCTGGAAGGATTTACG 1200
DB |||||||
QY 1141 GATCCATCTCTGGAAGGCTCTCAAAAAAGATTAAGAGATCTGGAAGGATTTACG 1200
DB |||||||
QY 1201 GTTTATCGGTTGTTCTCGAACCTTGGATGCAACACAGAAATTAACCTGGGAGATTTA 1260
DB |||||||
QY 1201 GTTTATCGGTTGTTCTCGAACCTTGGATGCAACACAGAAATTAACCTGGGAGATTTA 1260
DB |||||||
QY 1261 AAAATACAGATTTCTGGGCCCCACCTTGAATCATCGCTCCCAAGTATAGG 1320
DB |||||||
QY 1261 AAAATACAGATTTCTGGGCCCCACCTTGAATCATCGCTCCCAAGTATAGG 1320
DB |||||||
QY 1321 AGAAGAGAGAGCTCGTTGGGAGGAGAAAGAACCGTTCGAGTGAACAACAGAGTTCTG 1380
DB |||||||
QY 1321 AGAAGAGAGAGCTCGTTGGGAGGAGAAAGAACCGTTCGAGTGAACAACAGAGTTCTG 1380
DB |||||||
QY 1381 GCAAGACAAAGAACTATACCTCAATCAAGGAAAGAGCCGCGAGGAGACCATGCTGACA 1440
DB |||||||
QY 1381 GCAAGACAAAGAACTATACCTCAATCAAGGAAAGAGCCGCGAGGAGACCATGCTGACA 1440
DB |||||||
QY 1441 AGAATGAGGAGGAAAGAAAGAACTTCTTAAGCCTGTGCTGATGCACTCATTTAG 1500
DB |||||||
QY 1441 AGAATGAGGAGGAAAGAAAGAACTTCTTAAGCCTGTGCTGATGCACTCATTTAG 1500
DB |||||||
QY 1501 AAAACTACTACTCCAGTGACCCCATGATCAGGTGGGAGATTTCTACAGTGTGGAAACGA 1560
DB |||||||
QY 1501 AAAACTACTACTCCAGTGACCCCATGATCAGGTGGGAGATTTCTACAGTGTGGAAACGA 1560
DB |||||||
QY 1561 GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGAGGAGCAAGAAAGG 1620
DB |||||||
QY 1561 GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGAGGAGCAAGAAAGG 1620
DB |||||||
QY 1621 CCAAGCTGCTCGGCTCCGTGGGAACTCCAAAGACGAACTCGATTAAGATTTTAGAGA 1680
DB |||||||
QY 1621 CCAAGCTGCTCGGCTCCGTGGGAACTCCAAAGACGAACTCGATTAAGATTTTAGAGA 1680
DB |||||||

Db	3841	ACGCCATAACCAACGACCAAGAGGCGCTGTGGCGCTTCGCCACAAGACCTTGGCGG	3900
Qy	3901	ACCTTTTGAATCATTTATTGACGCGCTGTACACTGATAAGGATTTGGGAATATGTTCA	3960
Db	3901	ACCTTTTGGATCATTTATTGACGCGCTGTACACTGATAAGGATTTGGGAATATGTTCA	3960
Qy	3961	CTTTCATGAATGTCTGCTCTTTCCAGATGTGAAGAATTCATTTTGAATCAGGATTTGA	4020
Db	3961	CTTTCATGAATGTCTGCTCTTTCCAGATGTGAAGAATTCATTTTGAATCAGGATTTGA	4020
Qy	4021	ATGACCCCAATCCAGCTTCAGACGTGTGCTTGACACTTACGACAGAGGAGGAAGAGC	4080
Db	4021	ATGACCCCAATCCAGCTTCAGACGTGTGCTTGACACTTACGACAGAGGAGGAAGAGC	4080
Qy	4081	CAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGGAACCTACA	4140
Db	4081	CAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGGAACCTACA	4140
Qy	4141	CTGTGGCTGTTTATTTCAAGGGAGAAAGATAGCGCTGTGGAAAGGACCAAGTATTTCAGC	4200
Db	4141	CTGTGGCTGTTTATTTCAAGGGAGAAAGATAGCGCTGTGGAAAGGACCAAGTATTTCAGC	4200
Qy	4201	AAGCGGAATGGGAGCAGCAATGATGCGCTTGAATAATATATTTTCCCGAGATGCC	4260
Db	4201	AAGCGGAATGGGAGCAGCAATGATGCGCTTGAATAATATATTTTCCCGAGATGCC	4260
Qy	4261	ATCAGAGCGGTTTCATCGAACGGAAGTACAGACAGAGTAAAGAAATGAGGTGGGAAA	4320
Db	4261	ATCAGAGCGGTTTCATCGAACGGAAGTACAGACAGAGTAAAGAAATGAGGTGGGAAA	4320
Qy	4321	GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGGAGGCAT	4380
Db	4321	GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGGAGGCAT	4380
Qy	4381	GCAAGTGTGAGATTTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT	4440
Db	4381	GCAAGTGTGAGATTTTACTTGTCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT	4440
Qy	4441	AGTTTCTTCAGACAATGAACGAAGTGTCTCAATCGAAATAAATACAGAGTCNAATCG	4500
Db	4441	AGTTTCTTCAGACAATGAACGAAGTGTCTCAATCGAAATAAATACAGAGTCNAATCG	4500
Qy	4501	CTATTGTTGTTTAAATGATCTGTTTATAGCTGGATGCTTTTATTAACAAGTATTAGATT	4560
Db	4501	CTATTGTTGTTTAAATGATCTGTTTATAGCTGGATGCTTTTATTAACAAGTATTAGATT	4560
Qy	4561	TTTCTTCTATTAAACGGAATCTGACTTTGGTGAATGTGCAATTTCTCTTTTATTG	4620
Db	4561	TTTCTTCTATTAAACGGAATCTGACTTTGGTGAATGTGCAATTTCTCTTTTATTG	4620
Qy	4621	CTCTTTAAATAAATAATTCAGAGGATATTCTATGTGGAATAGATCTGTTTTCAT	4680
Db	4621	CTCTTTAAATAAATAATTCAGAGGATATTCTATGTGGAATAGATCTGTTTTCAT	4680
Qy	4681	CTGTGCTCCAGATTGTGACCCCTAGACTTTTCAATTGACAGTAAATAATGACTTTACTAG	4740
Db	4681	CTGTGCTCCAGATTGTGACCCCTAGACTTTTCAATTGACAGTAAATAATGACTTTACTAG	4740
Qy	4741	TAATAAAAAAAAAAAAAAAAAAAAAA	4764
Db	4741	TAATAAAAAAAAAAAAAAAAAAAAAA	4764
RESULT 8			
US-10-723-860-5003			
; Sequence 5003, Application US/10723860			
; Publication No. US20040253606A1			
; GENERAL INFORMATION:			
; APPLICANT: Aziz, Natasha			
; APPLICANT: Ginsburg, Wendy M.			
; APPLICANT: Zlotnik, Albert			
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &			
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators			

; FILE REFERENCE: 05882.0193.NPUS01			
; CURRENT APPLICATION NUMBER: US/10/723,860			
; CURRENT FILING DATE: 2003-11-26			
; PRIOR APPLICATION NUMBER: 60/429,739			
; PRIOR FILING DATE: 2002-11-26			
; NUMBER OF SEQ ID NOS: 8393			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 5003			
; LENGTH: 5425			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (86)..(87)			
; OTHER INFORMATION: n is a, c, g, or t			
US-10-723-860-5003			
Query Match 99.3%; Score 4732.4; DB 8; Length 5425;			
Best Local Similarity 99.7%; Pred. No. 0;			
Matches 4751; Conservative 0; Mismatches 11; Indels 2; Gaps 1;			
Qy	1	CTGTCTTGGTACCTGCGGTAGTAGCCTTGTCTGACGCGGATCTGCGCGCGCGAG	60
Db	101	CTGTCTTGGTACCTGCGGTAGTAGCCTTGTCTGACGCGGATCTGCGCGCGCGAG	160
Qy	61	AGCCTTTTATAGGTTGCTTTCCGGGGATGTGAAGATACAGAAATGACTGTGAATCAA	120
Db	161	AGCCTTTTATAGGTTGCTTTCCGGGGATGTGAAGATACAGAAATGACTGTGAATCAA	220
Qy	121	CCCATATCATCAGGAGCTGATTAATCTAGTGGAAAGATTAGAGTGTGCATCTTCACTA	180
Db	221	CCCATATCATCAGGAGCTGATTAATCTAGTGGAAAGATTAGAGTGTGCATCTTCACTA	280
Qy	181	TGATATGAGGAGCTCTCTGAGCTTATATTCTCTGTGGAAGATGTGACATATCCAGGCGGA	240
Db	281	TGATATGAGGAGCTCTCTGAGCTTATATTCTCTGTGGAAGATGTGACATATCCAGGCGGA	340
Qy	241	ACATCATGATGACGGGAAACACATGTACAGATGTCTTCCACCCGGGACGAGGGGTC	300
Db	341	ACATCATGATGACGGGAAACACATGTACAGATGTCTTCCACCCGGGACGAGGGGTC	400
Qy	301	CCCGAGACGAGGAGACATGGAGCCAGACCTCTCAGCACCATCTCTTTAGGCCCCCAAAATC	360
Db	401	CCCGAGACGAGGAGACATGGAGCCAGACCTCTCAGCACCATCTCTTTAGGCCCCCAAAATC	460
Qy	361	TGAGGCTGCTTCAACCTCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC	420
Db	461	TGAGGCTGCTTCAACCTCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC	520
Qy	421	CTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACCACTTTTG	480
Db	521	CTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACCACTTTTG	580
Qy	481	TACCTTCCCCCACCACCATCTCCGTCCAGCGAAGCCCTCTTCCCCCTGCCCCCAATCA	540
Db	581	TACCTTCCCCCACCACCATCTCCGTCCAGCGAAGCCCTCTTCCCCCTGCCCCCAATCA	640
Qy	541	GGCGCGCTTTCCCCCAACCAACAGATGAGGACCCCTTCCAGATTCCTCTGTTTCTCTC	600
Db	641	GGCGCGCTTTCCCCCAACCAACAGATGAGGACCCCTTCCAGATTCCTCTGTTTCTCTC	700
Qy	601	CCATGCCACCACTTGTCTTAATAAACCCCGAGTCCCTCCAGTCCCTGCGGACCTCTCGAC	660
Db	701	CCATGCCACCACTTGTCTTAATAAACCCCGAGTCCCTCCAGTCCCTGCGGACCTCTCGAC	760
Qy	661	AAGGACCTTCCCTTTCATGATGCCCTCCCTCCATGCTCCATCCCTCCCGCCCTCTCCAG	720
Db	761	AAGGACCTTCCCTTTCATGATGCCCTCCCTCCATGCTCCATCCCTCCCGCCCTCTCCAG	820
Qy	721	TCATGCCGACGAGGTTAAATTTATCAGTACCTCTCGGGCTATTCTCAACAATCTTCCAC	780
Db	821	TCATGCCGACGAGGTTAAATTTATCAGTACCTCTCGGGCTATTCTCAACAATCTTCCAC	880

781 QY CTCCAGTCTTAAAGTCTTCCAGAACCAACCTAGTCTTTCTCTGCCAGTCTTAATAACA 840
881 Db CTCCAGTCTTAAAGTCTTCCAGAACCAACCTAGTCTTTCTCTGCCAGTCTTAATAACA 940
841 QY GCAGTAGTCTCTATTTCCAGACATCTCCCTCCATACCACTCCCAAGGCTCCCAAGTCAGA 900
941 Db GCAGTAGTCTCTATTTCCAGACATCTCCCTCCATACCACTCCCAAGGCTCCCAAGTCAGA 1000
901 QY GAAGGTCTCCAGAAAGGCTGAAACACTATGATGACCAAGGACCGAGACACAGTCAATG 960
1001 Db GAAGGTCTCCAGAAAGGCTGAAACACTATGATGACCAAGGACCGAGACACAGTCAATG 1060
961 QY GCGAGGTGAGAGGACATCGGTCTCCCTGATCGCGGGAGCGAGGCGCGAGTCCCGACAGGA 1020
1061 Db GCGAGGTGAGAGGACATCGGTCTCCCTGATCGCGGGAGCGAGGCGCGAGTCCCGACAGGA 1120
1021 QY GAAGACAAGACAGCGCGGTACAGATCTGATTAAGACCGAGGGAGAAACCACTCTCGCCACC 1080
1121 Db GAAGACAAGACAGCGCGGTACAGATCTGATTAAGACCGAGGGAGAAACCACTCTCGCCACC 1180
1081 QY GCAGCTACGAAACGAGACAGGAGGAGAAACGAGGAGAGACACAGGATCGAGACAAACGAA 1140
1181 Db GCAGCTACGAAACGAGACAGGAGGAGAAACGAGGAGAGACACAGGATCGAGACAAACGAA 1240
1141 QY GATCACCATCTCTGGAAGGTCTTACAAAGAGATATAGAGATCTGGAAGGAGTTACG 1200
1241 Db GATCACCATCTCTGGAAGGTCTTACAAAGAGATATAGAGATCTGGAAGGAGTTACG 1300
1201 QY GTTTATCGGTGTCTCTGAAACCTGCTGGAATGACACACAGAAATACCTGGGGAGATTAATA 1260
1301 Db GTTTATCGGTGTCTCTGAAACCTGCTGGAATGACACACAGAAATACCTGGGGAGATTAATA 1360
1261 QY AAAATACAGATTTCTGGGCCCCACCTCGAGATTTGTAATCATGCTCCCAAGTAGGG 1320
1361 Db AAAATACAGATTTCTGGGCCCCACCTCGAGATTTGTAATCATGCTCCCAAGTAGGG 1420
1321 QY AGAAGAGAGAGCTCGTTGGGAGGAGAAAGACCGTTGGAGTGACACAGGATCTCG 1380
1421 Db AGAAGAGAGAGCTCGTTGGGAGGAGAAAGACCGTTGGAGTGACACAGGATCTCG 1480
1381 QY GCAAGACAAGAACTATACCTCAATCAAGGAAAGAGCCGAGGAGACCATCGCTGACA 1440
1481 Db GCAAGACAAGAACTATACCTCAATCAAGGAAAGAGCCGAGGAGACCATCGCTGACA 1540
1441 QY AGAATGAGGAGGAGAAAGAACTTTCTTAAGCCTGTGTGATTCGATGCACTCATTCAG 1500
1541 Db AGAATGAGGAGGAGAAAGAACTTTCTTAAGCCTGTGTGATTCGATGCACTCATTCAG 1600
1501 QY AAAACTACTCTCAGTGACCCCATGGATCAGGTGGGAGATTCATCAGTGGTTGGAAAGA 1560
1601 Db AAAACTACTCTCAGTGACCCCATGGATCAGGTGGGAGATTCATCAGTGGTTGGAAAGA 1660
1561 QY GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG 1620
1661 Db GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGGAGTTGGGGAGCAGGCAAGAAAGG 1720
1621 QY CCAAAGCTGCTCGGCTCCGTTGGGAAACCTTCAAAGACGAAGCTCGATGAAGATTTAGAGA 1680
1721 Db CCAAAGCTGCTCGGCTCCGTTGGGAAACCTTCAAAGACGAAGCTCGATGAAGATTTAGAGA 1780
1681 QY GTTCCAGTGAATCCAGTAGTGAGTCTGATGAGACACGACCTGTTCTAGCAGTCAAGCT 1740
1781 Db GTTCCAGTGAATCCAGTAGTGAGTCTGATGAGACACGACCTGTTCTAGCAGTCAAGCT 1840
1741 QY CTGAAGTTTTTGACGTTATTTGACAGAAATCAAAACGCAAAAGGCCACCTGACCGACTTC 1800
1841 Db CTGAAGTTTTTGACGTTATTTGACAGAAATCAAAACGCAAAAGGCCACCTGACCGACTTC 1900
1801 QY ATGATGAATCTTGGTACACGATTCAGGGCAGATGAATGATGACCACTCTGCAAAATGCA 1860
1901 Db ATGATGAATCTTGGTACACGATTCAGGGCAGATGAATGATGACCACTCTGCAAAATGCA 1958
1861 QY GCGCAAGGCAAGACGACAGGAAATTAGGCACAGCATTTATCTCTGAGAGAGGCCATCA 1920

1959 Db GCGCAAGGCAAGACGACAGGAATTAGGCACAGCATTTATCTTGGAGAGAGGCCATCA 2018
1921 QY AGCCCTGCTGCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGATCACAGTCT 1980
2019 Db AGCCCTGCTGCTCTATGACCAACAAATGCTGGCAGACTTTTCCACTACCGATCACAGTCT 2078
1981 QY CCGCGCTCTAGAACTTTTAACTGACAGGCGCACTGTTATAGAAATACATGATCACAGT 2040
2079 Db CCGCGCTCTAGAACTTTTAACTGACAGGCGCACTGTTATAGAAATACATGATCACAGT 2138
2041 QY ATATCTTTGAAGGATTTCTATGTTTGCACATGCCCCCTGACCAATATTTCCACTGTGTA 2100
2139 Db ATATCTTTGAAGGATTTCTATGTTTGCACATGCCCCCTGACCAATATTTCCACTGTGTA 2198
2101 QY AAGTAAATAGATTAACATAGACTACAGATTTCAATTTCAATGAAGAGATGATCGCGAGA 2160
2199 Db AAGTAAATAGATTAACATAGACTACAGATTTCAATTTCAATGAAGAGATGATCGCGAGA 2258
2161 QY ATTTTGTGTGAAAGGCTTTGAACTCTTTTTCACCTGTTCTTATTCAGAGATATTTTGAAT 2220
2259 Db ATTTTGTGTGAAAGGCTTTGAACTCTTTTTCACCTGTTCTTATTCAGAGATATTTTGAAT 2318
2221 QY TATATGACTGGAATCTTAAAGGTCTTTGTTTGAAGACAGCCCTCTGCTGCCAAGAT 2280
2319 Db TATATGACTGGAATCTTAAAGGTCTTTGTTTGAAGACAGCCCTCTGCTGCCAAGAT 2378
2281 QY TTCAATTTTCATGCCACCGTTTGTAAAGATTTCTTCCAGATGAGGAAAGAAAGTGTGCTCA 2340
2379 Db TTCAATTTTCATGCCACCGTTTGTAAAGATTTCTTCCAGATGAGGAAAGAAAGTGTGCTCA 2438
2341 QY TGCAACAGATTTCTCTGTACTTGTAAAGGTGACAGAAAGCCCTGCTGAGGAGAGA 2400
2439 Db TGCAACAGATTTCTCTGTACTTGTAAAGGTGACAGAAAGCCCTGCTGAGGAGAGA 2498
2401 QY TTGCCAATATGCTTCAAGTGGAGGAGCTGGAGTGGCAGAAATATGCAAGAGATTCGAAG 2460
2499 Db TTGCCAATATGCTTCAAGTGGAGGAGCTGGAGTGGCAGAAATATGCAAGAGATTCGAAG 2558
2461 QY GCATGATTTGTATCAACCCCTGGGACGAAACCAAGCTCTGTCCTATCGATCAACTGGATC 2520
2559 Db GCATGATTTGTATCAACCCCTGGGACGAAACCAAGCTCTGTCCTATCGATCAACTGGATC 2618
2521 QY GTGAAACAGTTCAACCCCGATGTGATTACTTTTCCGATTAATCGTCCACTTTGGGATACGC 2580
2619 Db GTGAAACAGTTCAACCCCGATGTGATTACTTTTCCGATTAATCGTCCACTTTGGGATACGC 2678
2581 QY CTGCACAGTTGATTAAGCAGGAGACCCACAGTACCAAAACCTGTGGAAGAGTTATGTA 2640
2679 Db CTGCACAGTTGATTAAGCAGGAGACCCACAGTACCAAAACCTGTGGAAGAGTTATGTA 2738
2641 QY AACTTCCGCACTCTCTAGCAAAATAGTCCCAAGTCAAAACAACTGACAAACAGAAAGCTGG 2700
2739 Db AACTTCCGCACTCTCTAGCAAAATAGTCCCAAGTCAAAACAACTGACAAACAGAAAGCTGG 2798
2701 QY CACAGAGGAGGAGGAGCCCTCCAAAAATACGCGAGAAAGAAATCAATGAGACGAGAGATA 2760
2799 Db CACAGAGGAGGAGGAGGAGCCCTCCAAAAATACGCGAGAAAGAAATCAATGAGACGAGAGATA 2858
2761 QY CCGTGGAGCTAAGTAGCCAGGATTTCTGGAAACTGGCATCCGTTCTGATGTTCTGTCAGC 2820
2859 Db CCGTGGAGCTAAGTAGCCAGGATTTCTGGAAACTGGCATCCGTTCTGATGTTCTGTCAGC 2918
2821 QY ATGCAATGATGCTACTGTTCTGACCCATCATATCGCTACCAACCAATGCTAATGCATT 2880
2919 Db ATGCAATGATGCTACTGTTCTGACCCATCATATCGCTACCAACCAATGCTAATGCATT 2978
2881 QY TGGACAAGTTGATGAGATATCTTTCCAAAGATGTTGTCTGTTGACGCTGGCCATGACTC 2940
2979 Db TGGACAAGTTGATGAGATATCTTTCCAAAGATGTTGTCTGTTGACGCTGGCCATGACTC 3038
2941 QY ATCCAAAGTCAATCAATTTAAATTTTGGAAATGAAATCTGATCATGCGCAGAAATTCATTCTA 3000

OTHER INFORMATION: Putative ribonuclease III
US-10-205-331-117

Query Match 34.0%; Score 1619.6; DB 7; Length 1626;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1622; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	2745	ATGAGAGGAGGTAACGGTGGAGCTAAGTAGCCCAAGGATTCGGAATCTGGCAATCGGCATCCGT	2804
Db	1	ATGAGAGGAGGTAACGGTGGAGCTAAGTAGCCCAAGGATTCGGAATCTGGCAATCGGCATCCGT	60
Qy	2805	TCTGATCTCTGTCAGCATGCAATGATCTACTCTCTGACCCCATCATATCCGCTACCCAC	2864
Db	61	TCTGATCTCTGTCAGCATGCAATGATCTACTCTCTGACCCCATCATATCCGCTACCCAC	120
Qy	2865	CAATGCTTAATGCAATTTGGCAAGTTGATAGGATATCTTTTCAAGATCGTGTCTGTG	2924
Db	121	CAATGCTTAATGCAATTTGGCAAGTTGATAGGATATCTTTTCAAGATCGTGTCTGTG	180
Qy	2925	CAGTGGCCATGACTCATCCAGTCATCTTAATTTTGGATGAATCTTGATCATGCC	2984
Db	181	CAGTGGCCATGACTCATCCAGTCATCTTAATTTTGGATGAATCTTGATCATGCC	240
Qy	2985	AGGAATTCATTTCTAACTGTGGAAATTCGGCAGCCCAAAATACGGAGACAGAAAAGTTTCAT	3044
Db	241	AGGAATTCATTTCTAACTGTGGAAATTCGGCAGCCCAAAATACGGAGACAGAAAAGTTTCAT	300
Qy	3045	CACATGCATATGCGGAAGAAAGGATTAACACCTTGTATTAATATCATGTCAACGCTTGGC	3104
Db	301	CACATGCATATGCGGAAGAAAGGATTAACACCTTGTATTAATATCATGTCAACGCTTGGC	360
Qy	3105	CAAGATGACCAACTCCTCGAGGATTAACACCAATGAACGGTTGGAAATTCCTGGGTGAT	3164
Db	361	CAAGATGACCAACTCCTCGAGGATTAACACCAATGAACGGTTGGAAATTCCTGGGTGAT	420
Qy	3165	GCTGTGTTGCAATTTCTGACAGCGTCCATTTGTACTATTTTCTCTAGTCTGGAAGAA	3224
Db	421	GCTGTGTTGCAATTTCTGACAGCGTCCATTTGTACTATTTTCTCTAGTCTGGAAGAA	480
Qy	3225	GGAGGATTAGCAACCTATCGGACTGCCATTTGTCAGAAATCAGCACCTTGCCATGCTAGCA	3284
Db	481	GGAGGATTAGCAACCTATCGGACTGCCATTTGTCAGAAATCAGCACCTTGCCATGCTAGCA	540
Qy	3285	AAGAAATTCGAATGATGATCAATTTATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAA	3344
Db	541	AAGAAATTCGAATGATGATCAATTTATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAA	600
Qy	3345	TCGGACCTTCGACATGCAATGGCCAAATTTGTTGAAAGCGTTAATAGGAGCTGTTTACTTG	3404
Db	601	TCGGACCTTCGACATGCAATGGCCAAATTTGTTGAAAGCGTTAATAGGAGCTGTTTACTTG	660
Qy	3405	GAGGGAAGCCTGGAGGAAGCAAGCAGTTAATTTGGACGCTTGCTCTTTAATGATCCGGAC	3464
Db	661	GAGGGAAGCCTGGAGGAAGCAAGCAGTTAATTTGGACGCTTGCTCTTTAATGATCCGGAC	720
Qy	3465	CTGGCGAAGTCTGGCTCAATTAATCTCTCCACCACTCAACTACAGAGCAAAATACT	3524
Db	721	CTGGCGAAGTCTGGCTCAATTAATCTCTCCACCACTCAACTACAGAGCAAAATACT	780
Qy	3525	GATCGACAACTTAATGAACTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCA	3584
Db	781	GATCGACAACTTAATGAACTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCA	840
Qy	3585	ATTGGAGTAAATTTTACTCATGTTTCGACTTCTGGCAAGGCAATTCACATTTGAGAACTGTG	3644
Db	841	ATTGGAGTAAATTTTACTCATGTTTCGACTTCTGGCAAGGCAATTCACATTTGAGAACTGTG	900
Qy	3645	GGATTTAAACATCTGACCCCTAGGCCCAATCAGAGAAATGAAATTCCTAGTGTACTCCATA	3704
Db	901	GGATTTAAACATCTGACCCCTAGGCCCAATCAGAGAAATGAAATTCCTAGTGTACTCCATA	960
Qy	3705	ATGCAACTGTGAGCCACAGAGTACTTATTCATTTCCAGATCATCATGAGGACAC	3764

Db	961	ATGCAACTGTGTAGCCACAGAGTACTTATTATTATTTCCAGATCATCATGAGGACAC	1020
Qy	3765	TTAACTTTTGTGGGAAGCTCTTTGGTGAATAAATAAGAACTCAGGCCAAGGTAGCGAGGAG	3824
Db	1021	TTAACTTTTGTGGGAAGCTCTTTGGTGAATAAATAAGAACTCAGGCCAAGGTAGCGAGGAG	1080
Qy	3825	CTGGGCATGACGAGTAGCCATCAACCAACGACAAAGACCAAGAGGCCCTGTGGCGCTTCGC	3884
Db	1081	CTGGGCATGACGAGTAGCCATCAACCAACGACAAAGACCAAGAGGCCCTGTGGCGCTTCGC	1140
Qy	3885	ACCAAGACCTTTGGCGGACCTTTTGGGAATCATTTATTTCAGCGCTGTACACTGATAAGGAT	3944
Db	1141	ACCAAGACCTTTGGCGGACCTTTTGGGAATCATTTATTTCAGCGCTGTACACTGATAAGGAT	1200
Qy	3945	TTGGAATATGTTCTATCTTCACTGAATGTCTGCTCTTTTCCACGATTTGAAAGAAATTCATT	4004
Db	1201	TTGGAATATGTTCTATCTTCACTGAATGTCTGCTCTTTTCCACGATTTGAAAGAAATTCATT	1260
Qy	4005	TTGGAATCAGGATTTGGAATGACCCCAATCCAGGCTTCAGCAGTGTCTGCTTGCACACTTAGG	4064
Db	1261	TTGGAATCAGGATTTGGAATGACCCCAATCCAGGCTTCAGCAGTGTCTGCTTGCACACTTAGG	1320
Qy	4065	ACAGAAGAAAAGAGCCAGACATTTCTCTGTACAAGACTCTGCAGACAGTGGGCCCCATCC	4124
Db	1321	ACAGAAGAAAAGAGCCAGACATTTCTCTGTACAAGACTCTGCAGACAGTGGGCCCCATCC	1380
Qy	4125	CATGCCCGAACTCCTACACTGTGGCTGTTTATTTTCAAGGAGAAAGAAATAGGCTGTGGAAA	4184
Db	1381	CATGCCCGAACTCCTACACTGTGGCTGTTTATTTTCAAGGAGAAAGAAATAGGCTGTGGAAA	1440
Qy	4185	GGACCAAGTATTCAGCAAGCGGAAATGGGAGCAGCAATGCGCTTGAATAATATTAAT	4244
Db	1441	GGACCAAGTATTCAGCAAGCGGAAATGGGAGCAGCAATGCGCTTGAATAATATTAAT	1500
Qy	4245	TTTCCCGCAGATGGCCCATCAGAAAGCGTTTCATCGAACGGAAGTACAGACAGAGTTAAAA	4304
Db	1501	TTTCCCGCAGATGGCCCATCAGAAAGCGTTTCATCGAACGGAAGTACAGACAGAGTTAAAA	1560
Qy	4305	GAAATGAGTGGGAAAAGAGAGCATCAAGAGAGAGCCAGATGAGACTGAAAGACATCAAG	4364
Db	1561	GAAATGAGTGGGAAAAGAGAGCATCAAGAGAGAGCCAGATGAGACTGAAAGACATCAAG	1620
Qy	4365	AAATAA 4370	
Db	1621	AAATAA 1626	

RESULT 10
US-10-103-313-23
; Sequence 23, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ07C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-103-313-23

Query Match 29.4%; Score 1398.6; DB 5; Length 1458;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1412; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
Qy 3234 GCAACCTATCGGACTGCCATTGTTTCAAGATTCAGACCTTGCCTATGCTAGCAAGAACTT 3293
Db 24 GCAACCTATCGGACTGCCATTGTTTCAAGATTCAGACCTTGCCTATGCTAGCAAGAACTT 83

QY	3294	GAACATGATCCCAATTTATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAAATCGGACCTT	3353
Db	84	GAACATGATCCCAATTTATGCTGTATGCTCAGCGGCTGACCTTTGTAGAGAAATCGGACCTT	143
QY	3354	CGACATGCAATGGCCAAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGAAGC	3413
Db	144	CGACATGCAATGGCCAAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGAAGC	203
QY	3414	CTGGAGGAGCCCAAGCAGTTATTTGGACGCTTGCTCTTTAATGATCCGACCTGCGCGAA	3473
Db	204	CTGGAGGAGCCCAAGCAGTTATTTGGACGCTTGCTCTTTAATGATCCGACCTGCGCGAA	263
QY	3474	GTCTGGCTCAATTTATCTCTCCACCACTCCAACTACAAGAGCCAAATACTGATCGACAA	3533
Db	264	GTCTGGCTCAATTTATCTCTCCACCACTCCAACTACAAGAGCCAAATACTGATCGACAA	323
QY	3534	CTTATTTGAAACTTCTCAGTTCTACAAAATCTTACTGAGTTTCAAGAGCAATTTGGAGTA	3593
Db	324	CTTATTTGAAACTTCTCAGTTCTACAAAATCTTACTGAGTTTCAAGAGCAATTTGGAGTA	383
QY	3594	ATTTTTTACTCATGTTCTGACCTTCTGGCAAGGCGATTACATTTGAGAACTGTGGGATTTAAC	3653
Db	384	ATTTTTTACTCATG-TGACCTTCTGGCAAGGCGATTACATTTGAGAACTGTGGGATTTAAC	442
QY	3654	CATCTGACCCCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTG	3713
Db	443	CATCTGACCCCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTG	502
QY	3714	GTAGCCACAGAGTACTTATTCATTTCTCCAGATCATCATGAGGACACTTAACTTTG	3773
Db	503	GTAGCCACAGAGTACTTATTCATTTCTCCAGATCATCATGAGGACACTTAACTTTG	562
QY	3774	TTGGGAGCTCTTTGGTGAATATAGAACTCAGSCCAAGGTAGCGGAGGAGCTGGGCATG	3833
Db	563	TTGGGAGCTCTTTGGTGAATATAGAACTCAGSCCAAGGTAGCGGAGGAGCTGGGCATG	622
QY	3834	CAGGAGTACGCCATAACCAACGACGACCAAGAGGCGCTGTGGCGCTTCGCACCAAGACC	3893
Db	623	CAGGAGTACGCCATAACCAACGACGACCAAGAGGCGCTGTGGCGCTTCGCACCAAGACC	682
QY	3894	TTGGCGGACCTTTTGGAAATCAATTTATTTGAGCGCTGTACACTGATAGGATTTGGAATAT	3953
Db	683	TTGGCGGACCTTTTGGAAATCAATTTATTTGAGCGCTGTACACTGATAGGATTTGGAATAT	742
QY	3954	GTTTCATCTTTTCATGAATGTCTGCTCTTTCCAGATTGAAAGAAATTCATTTTGAATCAG	4013
Db	743	GTTTCATCTTTTCATGAATGTCTGCTCTTTCCAGATTGAAAGAAATTCATTTTGAATCAG	802
QY	4014	GATTTGGAATGACCCCAATCCAGCTTCAGCAGTGTGCTTTGACACTTTAGGACAGAAGGA	4073
Db	803	GATTTGGAATGACCCCAATCCAGCTTCAGCAGTGTGCTTTGACACTTTAGGACAGAAGGA	862
QY	4074	AAAGAGCAGACATTCCTCTGTACAAGACTCTGCAGAGTGGGCCCATCCATGCCCGA	4133
Db	863	AAAGAGCAGACATTCCTCTGTACAAGACTCTGCAGAGTGGGCCCATCCATGCCCGA	922
QY	4134	ACCTACACTGTGGCTGTTTATTTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGACCAAGT	4193
Db	923	ACCTACACTGTGGCTGTTTATTTTCAAGGGAGAAAGATAGGCTGTGGGAAAGGACCAAGT	982
QY	4194	ATTGAGCAAGCGGAAATGGGAGCAGCAATGATGCTGTGAAAAATATAATTTTCCCGAG	4253
Db	983	ATTGAGCAAGCGGAAATGGGAGCAGCAATGATGCTGTGAAAAATATAATTTTCCCGAG	1042
QY	4254	ATGCCCCATCAGAAAGCGTTTCATCGAAGGAGTACAGACAGAGTTTAAAGAAATGAGG	4313
Db	1043	ATGCCCCATCAGAAAGCGTTTCATCGAAGGAGTACAGACAGAGTTTAAAGAAATGAGG	1102
QY	4314	TGGGAAAGAGGACATCAAGAGAGAGGACCATGATGAGCTGAAAGCATCAAGAAATAAAGG	4373
Db	1103	TGGGAAAGAGGACATCAAGAGAGAGGACCATGATGAGCTGAAAGCATCAAGAAATAAAGG	1162

QY	4374	AGGGCATGCAAGTGTGGAGTATTTACTTCTCAGTAACCTGTGACTGTGTCTATTATGAGAC	4433
Db	1163	AGGGCATGCAAGTGTGGAGTATTTACTTCTCAGTAACCTGTGACTGTGTCTATTATGAGAC	1222
QY	4434	CTAGCCTAGTCTTTTCTGTCAGACAATGAAAGTGTGCTCATTTGAAATAAATAACAGAGT	4493
Db	1223	CTAGCCTAGTCTTTTCTGTCAGACAATGAAAGTGTGCTCATTTGAAATAAATAACAGAGT	1282
QY	4494	CAAAATCGCTATTTGTTGTTTTTAATGATCTGTTTTTACTGCGATGGTCTTTATTACAAAGTA	4553
Db	1283	CAAAATCGCTATTTGTTGTTTTTAATGATCTGTTTTTACTGCGATGGTCTTTATTACAAAGTA	1342
QY	4554	TTAGATTTTCTCTCTATTTAAACGGAACCTTGACTTTTGGTGAATGTGCAATTTCTCTTT	4613
Db	1343	TTAGATTTTCTCTCTATTTAAACGGAACCTTGACTTTTGGTGAATGTGCAATTTCTCTTT	1402
QY	4614	TATTTTGTCTCTTTTAAATAATAAATTTCAAGAGCATA	4650
Db	1403	TATTTTGTCTCTTTTAAATAATAAATTTCAAGAGCATA	1439
RESULT 11			
US-10-103-313-187			
; Sequence 187, Application US/10103313			
; Publication No. US20030082758A1			
; GENERAL INFORMATION:			
; APPLICANT: Rosen et al.			
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies			
; FILE REFERENCE: PJZ07C1			
; CURRENT APPLICATION NUMBER: US/10/103,313			
; CURRENT FILING DATE: 2002-03-12			
; NUMBER OF SEQ ID NOS: 653			
; Prior Application removed - See File Wrapper or Palm			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 187			
; LENGTH: 1314			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-103-313-187			
Query Match 26.5%; Score 1262.4; DB 5; Length 1314;			
Best Local Similarity 99.4%; Pred. No. 0;			
Matches 1263; Conservative 3; Mismatches 4; Indels 0; Gaps 0;			
QY	3234	GCAACCTATCGGACTGCCATTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGC	3293
Db	2	GCAACCTATCGGACTGCCATTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGC	61
QY	3294	GAACTGGATCCATTTATGCTGTATGCTCAGGGGCTGACCTTTGTAGAGAAATCGGACCTT	3353
Db	62	GAACTGGATCCATTTATGCTGTATGCTCAGGGGCTGACCTTTGTAGAGAAATCGGACCTT	121
QY	3354	CGACATGCAATGGCCAAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGC	3413
Db	122	CGACATGCAATGGCCAAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGC	181
QY	3414	CTGAGAGAGCCAAAGCAGTTATTTGGAGCGTGTCTCTTTAATGATCCGACCTCGCGAA	3473
Db	182	CTGAGAGAGCCAAAGCAGTTATTTGGAGCGTGTCTCTTTAATGATCCGACCTCGCGAA	241
QY	3474	GTCTGGCTCAATTTATCTCTCCACCACTCCAACTACAAGAGCCAAATFACTGATCGACAA	3533
Db	242	GTCTGGCTCAATTTATCTCTCTCCACCACTCCAACTACAAGAGCCAAATFACTGATCGACAA	301
QY	3534	CTTATTTGAAACTTCTCCAGTTTCTACAAAATCTTACTGAGTTTGAAGAGCAATTTGGAGTA	3593
Db	302	CTTATTTGAAACTTCTCCAGTTTCTACAAAATCTTACTGAGTTTGAAGAGCAATTTGGAGTA	361
QY	3594	ATTTTACTCATGTTTGAAGCGTTTCAAGAGGCGATTCATTTGAGAACTGTGCGATTTAAC	3653
Db	362	ATTTTACTCATGTTTGAAGCGTTTCAAGAGGCGATTCATTTGAGAACTGTGCGATTTAAC	421
QY	3654	CATCTGACCCCTAGGCCCAATCAGAGAAATGGAATTTCTTAGTGTACTTCCATTAATGCAACTG	3713

```
Db 422 CATCTGACCTTAGCCACATCAGAGATGGAATTCCTAGGTACTCCATTAATGCAACTG 481
QY 3714 GTAGCCACAGAGTACTTATTCATTCATTTCCAGATCATCATGAAGACACTTAACTTTG 3773
Db 482 GTAGCCACAGAGTACTTATTCATTCATTTCCAGATCATCATGAAGACACTTAACTTTG 541
QY 3774 TTGCGAAGCTCTTTGGTGAATATAGAACTCAGGCCAAGGTAGCGGAGCTGGGCGATG 3833
Db 542 TTGCGAAGCTCTTTGGTGAATATAGAACTCAGGCCAAGGTAGCGGAGCTGGGCGATG 601
QY 3834 CAGAGTACGCCATAACCAACGACAGACCAAGAGGCTGTGGCGCTTCGCACCAAGACC 3893
Db 602 CAGAGTACGCCATAACCAACGACAGACCAAGAGGCTGTGGCGCTTCGCACCAAGACC 661
QY 3894 TTGCGCGACCTTTTGGAACTATTTATTTAGAGCGCTGTACACTCATTAAGATTTGGAATAT 3953
Db 662 TTGCGCGACCTTTTGGAACTATTTATTTAGAGCGCTGTACACTCATTAAGATTTGGAATAT 721
QY 3954 GTTCATACCTTCATGAATGCTGCTCTTTTCCAGGATTCAGGATTCATTTGAATCAG 4013
Db 722 GTTCATACCTTCATGAATGCTGCTCTTTTCCAGGATTCAGGATTCATTTGAATCAG 781
QY 4014 GATTGAATGACCCCAATCCCAAGCTTCAGAGTGTGTGACACTTAGGACAGAGGA 4073
Db 782 GATTGAATGACCCCAATCCCAAGCTTCAGAGTGTGTGACACTTAGGACAGAGGA 841
QY 4074 AAAGAGCCAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGA 4133
Db 842 AAAGAGCCAGACATTCCTCTGTACAAGACTCTGCAGACAGTGGGCCCATCCCATGCCGA 901
QY 4134 ACCTACACTGTGGCTGTTTATTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGACCAAGT 4193
Db 902 ACCTACACTGTGGCTGTTTATTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGACCAAGT 961
QY 4194 ATTTCAGCAAGCGGAATGGAGCAGCAATGGATGGCTGTGAAATAATATTTTCCCCAG 4253
Db 962 ATTTCAGCAAGCGGAATGGAGCAGCAATGGATGGCTGTGAAATAATATTTTCCCCAG 1021
QY 4254 ATGGCCCATCAGAAGCGGTTCATCGAAGGAAGTACAGACAGAGTTAAAGAAATGAGG 4313
Db 1022 ATGGCCCATCAGAAGCGGTTCATCGAAGGAAGTACAGACAGAGTTAAAGAAATGAGG 1081
QY 4314 TGGGAAGAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGG 4373
Db 1082 TGGGAAGAGAGCATCAAGAGAGAGCCAGATGAGACTGAAGACATCAAGAAATAAAGG 1141
QY 4374 AGGGCATGCAAGTGTGGAGTATTTACTCTCAGTAACTGCTGCTCTATTGAGAC 4433
Db 1142 AGGGCATGCAAGTGTGGAGTATTTACTCTCAGTAACTGCTGCTCTATTGAGAC 1201
QY 4434 CTAGCCTAGTTTCTTCGACAGCAATGAAGAAAGTGTGCTCATTTGAAATAAATAACAGAGT 4493
Db 1202 CTAGCCTAGTTTCTTCGACAGCAATGAAGAAAGTGTGCTCATTTGAAATAAATAACAGAGT 1261
QY 4494 CAATCGCTA 4503
Db 1262 CAATCGCAA 1271
```

RESULT 12

```
US-11-097-143-2408
; Sequence 2408, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
```

```
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2408
; LENGTH: 4275
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-2408
```

```
Query Match 16.2%; Score 770.2; DB 10; Length 4275;
Best Local Similarity 58.2%; Pred. No. 2e-208;
Matches 1489; Conservative 0; Mismatches 1023; Indels 45; Gaps 6;
```

```
QY 1758 ATTGCAGAAATCAAAACGCAAAAGGCCACCTGACCGACTTCATGATGAACCTTGTGATC 1817
Db 1151 ATGGAGGAGCTTTCGGCAAGGTGCAATCCGACGGTGTGCACGGGATCTCTGGCAC 1210
QY 1818 AACGATCCAGGCCAGATGAATGATGACCACTCTGCAAAATGCAAGCGCAAGGCAAGACGC 1877
Db 1211 AACGAGCGGTGAGATGAACGACGACCTCTTTGCGCTGTCTACGCAAGTCCCGGCGC 1270
QY 1878 ACAGGAATTAGGACACAGATTTATTCCTGAGAGAGGCCATCAAGCCCTGTCTCTATG 1937
Db 1271 ATTGGCATTCGACAGGCAATATCCGGCGAGACTGGCTGATTAAGTTGTGCGATCCAAAC 1330
QY 1938 ACCAAATATGCTGGCAGACTTTTCCACTACCGATCAGACTCTCCCGCTACGAACTTT 1997
Db 1331 ACACAAATGCAAGGACAGCTTTTCCACTACAGATCAGCATCTCACCGCCACTAATCTC 1390
QY 1998 TTAACGTGACAGGCCAACTGTTATAGAAATACGATGATCAGAGTATATCTTTGAAGGATTT 2057
Db 1391 CTGACAAAGACACCCACCATTAATCAAGCATGATGACACAGATTTCTATTCGAGGSGCTC 1450
QY 2058 TGTATGTTTGACATGCGCCCTGACCAATATTCACCTGTGTAAAGTAATTAAGATTCAAC 2117
Db 1451 TCACCTCTCTCGCATGTGCTCTCTCCGATCTGCGCGTCTGCAAGGTGATCGGCTTCAAC 1510
QY 2118 ATAGACTACAGATTCATTTTCATGAAGATGATGCGGAGAAATTTTGTGTGAAGGS 2177
Db 1511 ATCGAGTACACCATTTAGTACGAGAGGAGAAAGTGCAGAGAACTTCAACATCCACGAG 1570
QY 2178 CTGGAACCTCTTTCATCTGTTCTTATTCAGAGATATTTTGGAAATATATGACTGGAACTTT 2237
Db 1571 CTAGACATTTTTCAAATACCTGTTTATGAATCTGTAGAGCTGGTAGACTTTAATCTA 1630
QY 2238 AAAGTCTCTTTT---GTTTGAAGACAGCCCTCCCTGCTGCCCAAGATTTCAATTCATGCCA 2294
Db 1631 ATGCCCAACTTACCCTCGGAAACGTCGAGGAATCTGCGCAAGCTTTTCACTTCTTTCCG 1690
QY 2295 GGTGTTTGAAGATTTCTTCCAGATGAGAGGAAAGTGTCTGCTCCATGACACAGATTTCTC 2354
Db 1691 GGTGTTGCTCGGACCTGCGCAATTAATGGAAGAGGTTTGGCCATGGTCAGGATCTC 1750
QY 2355 CTGTACTTTGTTAAGGTGCGACAAAGCCCTGGTGTGCTGAGGAGGAGATTTGCCAATATGCTT 2414
Db 1751 CGCTACTTGTGGATAATTTCTGCACAGCTTTGTGGAACGCGACCACTACTGATCTTAAC 1810
QY 2415 CAGTGGGAGGAGCTGGAGTGGCAGAAATATGACAGAAATATGCAAGGCAATGATTTGTACC 2474
```


Db	1811	CAGATTAGTCAGAGCGAGTGGCAAACTACGTGGACTTCAAAAGGGAAATGCTGGTCACC	1870	Db	2891	AGTATATGGAGAAATCTGCGGAGCACCCGTTGCGAGAACAGGAGGCACCTTTGGGATCGC	2950
Qy	2475	AACCTTGGGACGAACCAAGCTCTGTCCGATCAATCAATGGATCGTGAAACAGTTCAA-	2533	Qy	3531	CAACTTATTTGAAACTTCTCCAGTTCTACAAAACTTACTAGTTTGAAGAGCAATTTGGA	3590
Db	1871	AAGCGGGTTTAAGCCGTGTTCGTACGCTTGACCAATTTGGACAGGAATAACTCCGAT	1930	Db	2951	AGCTGTATTTGATTTCTTAAAGGAGTTAAACCAAAATTCGAGGACTCTATTGGA	3010
Qy	2534	-----CCCCGATGTGATTAATTTTCGATATTCGTCACTTTGGG	2573	Qy	3591	GTAATTTTACTCATGTTTCGACTTCTGGCAAGGCACTTCAATTTGAGAACTTGTGGGATTT	3650
Db	1931	TTGCCCGAGTGGTAGATCGGAGACTGGAATCTCAATCCAGCAATCTGTCACCTTTGGC	1990	Db	3011	ATCAAGTTTAAAGCACTTCGGCTTTTGGCTCGCGCTTTTCCGATCGCTCCATTTGTTTC	3070
Qy	2574	ATACGCCCTTGCAAGTTATCGAGAGACCCACAGTACCAAAACTGTGGAGAGT	2633	Qy	3651	AACCATCTGACCTTAGGCCCAATCAGAGAAATGGAATTCCTAGTGTACTCCATAATGCA	3710
Db	1991	ATTTGTCATCTCAGTAAGCTACGTCTGGAATCCAGAGTACCAAGGCGTGGCGAG	2050	Db	3071	ACCACATTGACCTTTGGGTCAATCAGCGTTTAGAGTTCTTTGGGCGACACAGTGTGCG	3130
Qy	2634	TATGTGAAACTTCCGCACTCTAGCAATATGTCGCAAGTCAAACTGACAAAACAG	2693	Qy	3711	CTGGTAGCCACAGAGTACTTATTCATTTCCAGATCATCATGAAGGACACTTAAT	3770
Db	2051	TACGTTAAGTACCGTCATCTGATGGCAACATGTGCAAGCCCTCTTTCAAGGATAAGCG	2110	Db	3131	CTGATTTGCTCGGAGTACTATATCGTCACTTCCCTGAGCACCCACGAGGCGCACTTGTC	3190
Qy	2694	AGCTGGCACAGAGGAGGAGCCCTCCAAAATAACGCGAAGAAATACAAATGAGACGA	2753	Qy	3771	TTGTTGCGAAGCTCTTTTGGGAATCATTTATGCGCGCTGTACACTGATTAAGGATTTGGA	3830
Db	2111	AAGCTAGAGGAGAGGAGCAACGTCTTCAGGAGATGCGAACTCAGGGGCGCATGAACGA	2170	Db	3191	CTGCTACGATCTCTCATTTGGTCAATAATCGCACTCAAGCGGTGTTTGCATGATTTGGGA	3250
Qy	2754	GAAGTAAACGTTGGAGCTAAGTAGCAAGATTTCTGGAACCTGGCATCCGTTCTGATGTC	2813	Qy	3831	ATGCGAGGATACGCCATAACCAACGACAAAGACCAAGAGGCTGTGGCGCTTCGACCAAG	3890
Db	2171	ATATATCACAGTGGCGATCAGTCGAGGGCTTCTATCGCACCGGCATATGTCGCGAGTT	2230	Db	3251	ATGCCGAAATATGC-----CGTGTATGCCAATCCCAAGGCTGATTTGAAGACCAAA	3301
Qy	2814	TGTCAGCATGCAATGATGTACTGTTCTGACCCATCATATCCGTACCAACCAATGCCTA	2873	Qy	3891	ACCTTGGCGGACCTTTTGGGAATCATTTATGCGCGCTGTACACTGATTAAGGATTTGGA	3950
Db	2231	GTGAGCATGCCATGTTGATTCCTGCTTAACTTGGTCACTTCGCTTCAAGTCCGCTG	2290	Db	3302	GATCGTCCGATCTGCTGGAGCATTTCTCGGCGCTCTGTACGTGGAAGGCTCTCCTG	3361
Qy	2874	ATGCAATTTGCAAGTTGATGAGATATCTTCCAAAGATCGTTGTCTGTTGCACTGGCC	2933	Qy	3951	TATGTTCAATCTTTCAATGATGTCTCTTTCCACGATTTGAAGAAATTCATTTTGAAT	4010
Db	2291	GACCTGCTAGAGAGATATCGGTACCGCTTTAAAAATCGGTACTTCTCCAATTGGCG	2350	Db	3362	TATTTGAAACAGTTTGGCCATGATGTTTGTTCCTCCCGACTTCAGTTTGTATCATGAT	3421
Qy	2934	ATGACTCATCAAGTCATCAATTTAAATTTTGGATGATCTCTGATCATGCAAGAAATCA	2993	Qy	4011	CAGGATTTGGAATGACCCCAATCCCAAGCTTCCAGCTTTCAGCTGTTTGTGACACTTAGACA	4067
Db	2351	CTGAGCATCCCTCATCAAGGAGAACTACGGTACCAATCCGATACAGCCCGCTTAATCG	2410	Db	3422	CAGGACTGGAAACGATCCCAAGTCAAGCTGCAAGCTGCAAGTGTGCTCACAATTCGCACA	3481
Qy	2994	TTATCTAATCTGTGGAATTCGGAGCCCAATACGAGACAGAAAGTTTCATCATGATCAC	3053	Qy	4068	GAAGGAAAAGAGCGACGACATTCCTCTGTGTACAAGACTTGACAGACAGTGGGCCCCAT	4127
Db	2411	CTGACTAATCTCGGAATTCGTACCGGAGTACGAGATCGCAAGATCCATCATGAAC	2470	Db	3482	GATGGCGGCGAGCGGACATTCCTACTACAAGTGTGTGGAGGCCAGTGTGCTCAACTAAT	3541
Qy	3054	ATGCGGAAGAAAGGATTAACACCTTGTATTAATATCATGTACGCTTGGCCAGATGAC	3113	Qy	4128	GCCGGAACCTTACTCTGTGGCTGTTTATTTCAAGGAGAAAGAAATAGGCTGTGGGAAAAGGA	4187
Db	2471	ACAGCGAAGGGGTATCAACACATTAAGTGTGAGCATTTATGTCAGATTTGGCAAGGAT	2530	Db	3542	ACGCGGTATACAAGTGTGGCGCTTTATTTCCGCTCAAGCGGCTGGCCACTTCAAGTGGC	3601
Qy	3114	CCAACTCCCTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTT	3173	Qy	4188	CCAAGTATTCAGCAAGCGGAAATGGGAGCAGCAATGGAATGCGCTTGAATAAT-----AT	4241
Db	2531	GAGACTGTTTCGAACATAACCAATAGAGACTAGAGTTCTCGGTGATGCTGTAGTA	2590	Db	3602	TCCTCCATTCAGCAGGCGGAGATGAACGCCGCTAAGCAGGCGCTGGAGAAATTCAGGGAC	3661
Qy	3174	GAAATTTCTGACAGCGTCCATTTGTTACTATTTGTTTCTAGTCTGGAAGAAAGGAGATTA	3233	Qy	4242	AATTTTCCCAAGATGGCCCATCAGAAGCGGTTTCATCG	4278
Db	2591	GAAATTCCTCAGCTCAATCTGTTTATTTATGTTCCCTGAACTGGAGGAGGTTGTTG	2650	Db	3662	CTGTTTCCCACTGGATCACCAAAAGCGCGTATCG	3698
Qy	3234	GCAACTATCGAGTGCATTTGTCAGAAATCAGACACTTTCGCAATCTAGCAAAAGAACTT	3293				
Db	2651	GCACTTACCGAGCGCAATTTGTCGAACAGCACTTGGCTCTGTTGGCCAAAAGCTG	2710				
Qy	3294	GAACTGGATPCCATTTATGCTGATGCTCAGCGGCTGACCTTTGTAGAGAACTCGACCTT	3353				
Db	2711	CAACTGGAGAGTTTATGTCGTACGACCAACGATCCGATCTGTGCCACGAGTTGGAATG	2770				
Qy	3354	CGACATGCAATGCCCAATTTTGAAGGTTAATAAGGAGCTGTTTACTTGGAGGGAAGC	3413				
Db	2771	CGTACGCCATGCCAACTGTTTGAAGCCCTTAATGGCGCGCTTCTATTTGGATGGTGA	2830				
Qy	3414	CTGGAGGAACCAAGCAGTTATTTGGACGCTTCTCTTTA---ATGATCCGACCTGGCC	3470				
Db	2831	ATCAAGTGGCAGATGAGGTTTACGGATGCACTCTTCGGCAGGACGAGAACTGCTG	2890				
Qy	3471	GAAGTCTGGCTCAATTTATCTCTCCACCCACTCCAACCTACAGCAAAATCTGATCGA	3530				

RESULT 13
US-11-097-143-2407/c
; Sequence 2407, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932

Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20941
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 598, 632, 633
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-20941

Query Match 9.9%; Score 469.4; DB 8; Length 633;
Best Local Similarity 98.5%; Pred. No. 8.1e-123;
Matches 473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 58 GAGAGCCTTTTATAGGTTGCTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAAT 117
DB 152 GAGTTACTTTTCAGGTTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAAT 211
QY 118 CAACCCATATCATCAGGAGCTGATATCTAGTGGAGAGTTAGACGTGTGCATCTTCA 177
DB 212 CAACCCATATCATCAGGAGCTGATATCTAGTGGAGAGTTAGACGTGTGCATCTTCA 271
QY 178 CTATGATATCAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGC 237
DB 272 CTATGATATCAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGC 331
QY 238 GGAACATCATGTCAGGGAACACATGTCCAGAGATGTCTCCACCCGCGGAGGAGGC 297
DB 332 GGAACATCATGTCAGGGAACACATGTCCAGAGATGTCTCCACCCGCGGAGGAGGT 391
QY 298 GTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 357
DB 392 GTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 451
QY 358 ATCTGAGGCTGCTTCAACCTCAGAGGCTCTCTGTGCAATATCAATATGAACCTCCAAAGTG 417
DB 452 ATCTGAGGCTGCTTCAACCTCAGAGGCTCTCTGTGCAATATCAATATGAACCTCCAAAGTG 511
QY 418 CCGCTTCCACCACTTTCTCAAACTCTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 477
DB 512 CCGCTTCCACCACTTTCTCAAACTCTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 571
QY 478 TTGTACCCCTTCCCGCCACCCATGCTCCGTCAGGCGAAGGCGCTCTTCCCGCTGCGCCAA 537
DB 572 TTGTACCCCTTCCCGCCACCCATGCTCCGTCAGGCGAAGGCGCTCTTCCCGCTGCGCCAA 631

Search completed: December 27, 2005, 04:23:16
Job time : 3516 secs

This Page Blank (uspto)

Qy	61	AGCCTTTTATAGTGTCTTTCCGGGGATGGAAGATACAGAAATGATGTGAATCAA	120	Qy	1141	GATCACCATCTCTGGAAGGTCTCTCAAAAAAGAGTATAAGAGATCTGGAAGGAGTTACG	1200
Db	61	AGCCTTTTATAGTGTCTTTCCGGGGATGTGAAGATACAGAAATGACTGTGAATCAA	120	Db	1141	GATCACCATCTCTGGAAGGTCTCTCAAAAAAGAGTATAAGAGATCTGGAAGGAGTTACG	1200
Qy	121	CCCATATCATCAAGGAGCTGATTAATCTAGTGGAGAGTTAGAGGTGTGCATATCTCACTA	180	Qy	1201	GTTTATCGGTGTCTCTGAACTCTGTAACCTGTGGATGCACACAGAAATTACCTGGGGAGATTATTA	1260
Db	121	CCCATATCATCAAGGAGCTGATTAATCTAGTGGAGAGTTAGAGGTGTGCATATCTCACTA	180	Db	1201	GTTTATCGGTGTCTCTGAACTCTGTAACCTGTGGATGCACACAGAAATTACCTGGGGAGATTATTA	1260
Qy	181	TGATATGAGCGAGTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGCGGA	240	Qy	1261	AAAAATACAGATCTTTGGGCCCCACCCCTGGAGTTGTGAATCATCGCTCCCCAAGTAGGG	1320
Db	181	TGATATGAGCGAGTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGCGGA	240	Db	1261	AAAAATACAGATCTTTGGGCCCCACCCCTGGAGTTGTGAATCATCGCTCCCCAAGTAGGG	1320
Qy	241	ACATCATGATGACAGGAAAACAATGTCAAGAAATGTCTTCCACCCGGGACGAGGGGTC	300	Qy	1321	AGAAGAAGAGAGCTCTGTGGGAGGAAAAAGACCGTTGGAGTGAACAACACAGAGTTCTG	1380
Db	241	ACATCATGATGACAGGAAAACAATGTCAAGAAATGTCTTCCACCCGGGACGAGGGGTC	300	Db	1321	AGAAGAAGAGAGCTCTGTGGGAGGAAAAAGACCGTTGGAGTGAACAACACAGAGTTCTG	1380
Qy	301	CCGAGGACGAGGAGGACATGAGAGCCCTCAGGACCATCTTTAGGCCCCAAATC	360	Qy	1381	GCAAGAACAAGAACTATACCTCAATCAAGGAAAAAGAGCCCGAGGAGACCATGCTGTGACA	1440
Db	301	CCGAGGACGAGGAGGACATGAGAGCCCTCAGGACCATCTTTAGGCCCCCAATC	360	Db	1381	GCAAGAACAAGAACTATACCTCAATCAAGGAAAAAGAGCCCGAGGAGACCATGCTGTGACA	1440
Qy	361	TGAGGCTGTCTACCTTCAGCAGCCTCTGTGCAATATCAATGAACTTCCAAAGTGCCC	420	Qy	1441	AGATGAGGAGGAGAAAGAACTTCTTAAGCCTGTGTGATTCGATGCACCTCATTCAG	1500
Db	361	TGAGGCTGTCTACCTTCAGCAGCCTCTGTGCAATATCAATGAACTTCCAAAGTGCCC	420	Db	1441	AGATGAGGAGGAGAAAGAACTTCTTAAGCCTGTGTGATTCGATGCACCTCATTCAG	1500
Qy	421	CTTCCACACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACGAGCTTTG	480	Qy	1501	AAAACTACTCTCAGTGACCCCATCGATCAGGTGGAGATTCTACAGTGGTTGGAACGA	1560
Db	421	CTTCCACACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACGAGCTTTG	480	Db	1501	AAAACTACTCTCAGTGACCCCATCGATCAGGTGGAGATTCTACAGTGGTTGGAACGA	1560
Qy	481	TACCTTTCCCCCAACCATGCTCCGTGAGGGAAGGCCCTCTTCCGCCCTGCGCCATCA	540	Qy	1561	GTAGGCTTCTGTGATATATGACAAATTTGAGGAGGAGTTGGGAGCAGGCAAGAAAGG	1620
Db	481	TACCTTTCCCCCAACCATGCTCCGTGAGGGAAGGCCCTCTTCCGCCCTGCGCCATCA	540	Db	1561	GTAGGCTTCTGTGATATATGACAAATTTGAGGAGGAGTTGGGAGCAGGCAAGAAAGG	1620
Qy	541	GGCGCCTTTCCCAACCAAGATGAGGACCCCTTCCAGTTCTCTTGTGTTTCTC	600	Qy	1621	CCAAAGCTGTCTGGGCTCTCGTGGGAACCTCAAAAGAGGAGCTCGATGAGATTAGAGA	1680
Db	541	GGCGCCTTTCCCAACCAAGATGAGGACCCCTTCCAGTTCTCTTGTGTTTCTC	600	Db	1621	CCAAAGCTGTCTGGGCTCTCGTGGGAACCTCAAAAGAGGAGCTCGATGAGATTAGAGA	1680
Qy	601	CCATGCCACACCAATGCTTGTCTTAATAACCCCCAGTCCCTGGGGGCACTCTGGAC	660	Qy	1681	GTTCAGTGAATCCGAGTGTGAGTCTGTAGGACAGACACTGTCTTAGCAGCTCAGACT	1740
Db	601	CCATGCCACACCAATGCTTGTCTTAATAACCCCCAGTCCCTGGGGGCACTCTGGAC	660	Db	1681	GTTCAGTGAATCCGAGTGTGAGTCTGTAGGACAGACACTGTCTTAGCAGCTCAGACT	1740
Qy	661	AAGGCATTTTCCCTTCATGATGCCCTCCCTCCATGCTCATCCCGGCCCTCCAG	720	Qy	1741	CTGAAAGTTTGTGAGTTTATTCAGAAATCAAAACGCAAAAAAGGCCACCTCGACCTTC	1800
Db	661	AAGGCATTTTCCCTTCATGATGCCCTCCCTCCATGCTCATCCCGGCCCTCCAG	720	Db	1741	CTGAAAGTTTGTGAGTTTATTCAGAAATCAAAACGCAAAAAAGGCCACCTCGACCTTC	1800
Qy	721	TCATGCCGACAGGTTAAATTCAGTACCTCCGGGGTATTTCTCACCAAACTTCCAC	780	Qy	1801	ATGATGAATCTTTGGTACACGATCCAGGCCAGATGATGAGTGGACACTCTGCAAAATGCA	1860
Db	721	TCATGCCGACAGGTTAAATTCAGTACCTCCGGGGTATTTCTCACCAAACTTCCAC	780	Db	1801	ATGATGAATCTTTGGTACACGATCCAGGCCAGATGATGAGTGGACACTCTGCAAAATGCA	1860
Qy	781	CTCCAGTTTAAATAGTTTCCAGAACACCCCTAGTTCTTTCTGCCCAGTGTCTAATAA	840	Qy	1861	GCGCAAGGCAAGACGACAGGAAATAGGACAGCATTTATCTCTGGAGAGAGGCCATCA	1920
Db	781	CTCCAGTTTAAATAGTTTCCAGAACACCCCTAGTTCTTTCTGCCCAGTGTCTAATAA	840	Db	1861	GCGCAAGGCAAGACGACAGGAAATAGGACAGCATTTATCTCTGGAGAGAGGCCATCA	1920
Qy	841	GCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCACTCCCAAGGCTCCAGTGAGA	900	Qy	1921	AGCCCTGTCTCTATGACCAAACTGCTGGCAGACTTTTCCACTACCGGATCAAGTCT	1980
Db	841	GCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCACTCCCAAGGCTCCAGTGAGA	900	Db	1921	AGCCCTGTCTCTATGACCAAACTGCTGGCAGACTTTTCCACTACCGGATCAAGTCT	1980
Qy	901	GAAAGTCCCCAGAAAGGCTGAAAACATATGATGACCAAGGACCGGACCAAGTCAATG	960	Qy	1981	CCCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCACCGAGT	2040
Db	901	GAAAGTCCCCAGAAAGGCTGAAAACATATGATGACCAAGGACCGGACCAAGTCAATG	960	Db	1981	CCCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCACCGAGT	2040
Qy	961	GGCAGGTGAGAGGATCGCTCCCTGGATCGCGGGAGCGGCGCAGTCCCGACAGGA	1020	Qy	2041	ATATCTTTGAGGAGTTTCTATGTTTGCATGTCACATGCCCCCTGACCAATATTCACCTGTGTA	2100
Db	961	GGCAGGTGAGAGGATCGCTCCCTGGATCGCGGGAGCGGCGCAGTCCCGACAGGA	1020	Db	2041	ATATCTTTGAGGAGTTTCTATGTTTGCATGTTGCATGTCACATGCCCCCTGACCAATATTCACCTGTGTA	2100
Qy	1021	GAAAGACAGACAGCCGATACAGATCTGATTATGACCGAGGAGAAACAATCTCGCCACC	1080	Qy	2101	AAGTAATTAGAATCAACATAGACTACAGATTCTTTTCAATTTGAAAGAGATGATGCCGAGGA	2160
Db	1021	GAAAGACAGACAGCCGATACAGATCTGATTATGACCGAGGAGAAACAATCTCGCCACC	1080	Db	2101	AAGTAATTAGAATCAACATAGACTACAGATTCTTTTCAATTTGAAAGAGATGATGCCGAGGA	2160
Qy	1081	GCAGTACGACGAGAGCAGGAGCAACCGGAGAGACAGGCGATCGAGACACCGAA	1140	Qy	2161	ATTTTGTGTGAAAGGGCTTGAACTCTTTTCACTGTCTCTTATTCAGAGATATTTTGGAAAT	2220
Db	1081	GCAGTACGACGAGAGCAGGAGCAACCGGAGAGACAGGCGATCGAGACACCGAA	1140	Db	2161	ATTTTGTGTGAAAGGGCTTGAACTCTTTTCACTGTCTCTTATTCAGAGATATTTTGGAAAT	2220
				Qy	2221	TATATGACTGGAACTTTAAAGGTCTTTTGTGAAAGACAGGCCCTCCCTGTGCCCCAAGAT	2280

|||||
2221 TATATGACTGGAACTTAAAGTCTTTTGGTTGAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
QY
2281 TTCAATTCATGCCACGTTTGTGAAGATTTCTTCCAGATGGAGGAAGAAAGTGTGTCCA 2340
Db
2281 TTCAATTCATGCCACGTTTGTGAAGATTTCTTCCAGATGGAGGAAGAAAGTGTGTCCA 2340
QY
2341 TGCACCAAGATTCCTCTGTACTTGTGAAGTGCAGCAAGCCCTGGTCTGAGGAGGAGA 2400
Db
2341 TGCACCAAGATTCCTCTGTACTTGTGAAGTGCAGCAAGCCCTGGTCTGAGGAGGAGA 2400
QY
2401 TTGCCAATATATCTTCAGTGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAAAG 2460
Db
2401 TTGCCAATATATCTTCAGTGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAAAG 2460
QY
2461 GCATGATTTGTTACCAACCTCGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
Db
2461 GCATGATTTGTTACCAACCTCGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
QY
2521 GTGAACAGTTCAACCCCGATGTGATTAATTTTCCGATTAATCGTCCACTTTGGGATACGCC 2580
Db
2521 GTGAACAGTTCAACCCCGATGTGATTAATTTTCCGATTAATCGTCCACTTTGGGATACGCC 2580
QY
2581 CTGCACAGTTGAGTTATGACAGGAGACCCACAGTACCAAAACTGTGGAAGAGTTATGTGA 2640
Db
2581 CTGCACAGTTGAGTTATGACAGGAGACCCACAGTACCAAAACTGTGGAAGAGTTATGTGA 2640
QY
2641 AACTTCGCCACCTCTAGCAAAATAGTCCAAAGTCAAACTGACAAACAGAAAGCTGG 2700
Db
2641 AACTTCGCCACCTCTAGCAAAATAGTCCAAAGTCAAACTGACAAACAGAAAGCTGG 2700
QY
2701 CACAGAGGAGGAGAGCCCTCCAAAATACGGCAGAGAGATACAAATGAGACGAGAGTAA 2760
Db
2701 CACAGAGGAGGAGAGCCCTCCAAAATACGGCAGAGAGATACAAATGAGACGAGAGTAA 2760
QY
2761 CGGTGGAGCTAAGTAGCAAGGATTTCTGGAAGAACTGGCATCCGTTCTGATGCTGTGACG 2820
Db
2761 CGGTGGAGCTAAGTAGCAAGGATTTCTGGAAGAACTGGCATCCGTTCTGATGCTGTGACG 2820
QY
2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGTCACCAATGCCCTAATGCATT 2880
Db
2821 ATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGTCACCAATGCCCTAATGCATT 2880
QY
2881 TGGCAAGATTGATAGGATATACCTTCCAGATCGTTGTCTGTCAGTGCGCCATGATC 2940
Db
2881 TGGCAAGATTGATAGGATATACCTTCCAGATCGTTGTCTGTCAGTGCGCCATGATC 2940
QY
2941 ATCCAAAGTCATATTTAAATTTTGAAGTGAATCTGATCATGCGCAGGAATTCATTATCTA 3000
Db
2941 ATCCAAAGTCATATTTAAATTTTGAAGTGAATCTGATCATGCGCAGGAATTCATTATCTA 3000
QY
3001 ACTGTGGAATTCGGCAGCCCAAAATACGGGACAGAGAAAGTTTCATCATGCAATGCGGA 3060
Db
3001 ACTGTGGAATTCGGCAGCCCAAAATACGGGACAGAGAAAGTTTCATCATGCAATGCGGA 3060
QY
3061 AGAAAGGATTAACACCTTGATTAATATCATGTGTCAGCGCTTGGCCAAAGTGAACCACTC 3120
Db
3061 AGAAAGGATTAACACCTTGATTAATATCATGTGTCAGCGCTTGGCCAAAGTGAACCACTC 3120
QY
3121 CCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTTGTTGAATTC 3180
Db
3121 CCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGTGATGCTGTTGTTGAATTC 3180
QY
3181 TGACCAAGCTCCATTTGCTATTTGTTTCTTAGTCTGGAGAGAGGAGTTAGCAACCT 3240
Db
3181 TGACCAAGCTCCATTTGCTATTTGTTTCTTAGTCTGGAGAGAGGAGTTAGCAACCT 3240
QY
3241 ATCGGATGCGATTTGTTGAGATCAGCACTTGGCTGCTAGCAAGAACTTGAACCTGG 3300
Db
3241 ATCGGATGCGATTTGTTGAGATCAGCACTTGGCTGCTAGCAAGAACTTGAACCTGG 3300
QY
3301 ATCCATTTATGCTGTATGCTCACGGGCTGACCTTTGTTGATAGAAATCGGACCTTCGACATG 3360
|||||

Db
3301 ATCCATTTATGCTGTATGCTCACGGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
QY
3361 CAATGGCCAAATGTTTGAAGCGTTAAATAGAGCTGTTTAACTTTGGAGGAAAGCTGGAGG 3420
Db
3361 CAATGGCCAAATGTTTGAAGCGTTAAATAGAGCTGTTTAACTTTGGAGGAAAGCTGGAGG 3420
QY
3421 AAGCCAAAGCAGTTATTTGGACGCTTGTCTTTAATGATCCGACCTGCGGAGTCTGGC 3480
Db
3421 AAGCCAAAGCAGTTATTTGGACGCTTGTCTTTAATGATCCGACCTGCGGAGTCTGGC 3480
QY
3481 TCAATTAATCTCTCCACCCACTCCAACTACAAGAGCCAAATACATGATCGAACAACTTATG 3540
Db
3481 TCAATTAATCTCTCCACCCACTCCAACTACAAGAGCCAAATACATGATCGAACAACTTATG 3540
QY
3541 AAACTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATTTGAGTAAATTTTAA 3600
Db
3541 AAACTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATTTGAGTAAATTTTAA 3600
QY
3601 CTGATGTTGAGCTTCTGGCAGGCGATTCATCATTTGAGAACTGTGGGATTTAAACCATCTGA 3660
Db
3601 CTGATGTTGAGCTTCTGGCAGGCGATTCATCATTTGAGAACTGTGGGATTTAAACCATCTGA 3660
QY
3661 CCCTAGGCCCAATCAGAGAAATGGAATTTCTAGGTGACTCCATATGCAACTGGTAGCCA 3720
Db
3661 CCCTAGGCCCAATCAGAGAAATGGAATTTCTAGGTGACTCCATATGCAACTGGTAGCCA 3720
QY
3721 CAGAGTACTTATTCATTTCCAGATCATATGAAAGGACACTTAACTTTGTTGCGAA 3780
Db
3721 CAGAGTACTTATTCATTTCCAGATCATATGAAAGGACACTTAACTTTGTTGCGAA 3780
QY
3781 GCTCTTTGTTGGAATTAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGATGAGGAGT 3840
Db
3781 GCTCTTTGTTGGAATTAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGATGAGGAGT 3840
QY
3841 ACGCCATAAACCAACGACCAAGAGGCGCTGTGGCGCTTCGCAACCAAGACCTTTGGCGG 3900
Db
3841 ACGCCATAAACCAACGACCAAGAGGCGCTGTGGCGCTTCGCAACCAAGACCTTTGGCGG 3900
QY
3901 ACCTTTTGGAAATCAATTTTTCAGCGCTGTACACTGATAGGAATTTGGAATATGTTTCA 3960
Db
3901 ACCTTTTGGAAATCAATTTTTCAGCGCTGTACACTGATAGGAATTTGGAATATGTTTCA 3960
QY
3961 CTCTTCAATGATGCTGCTTCTTTTCCAGATTCAGAAAGATTTCAATTTGATCAGGATTTGA 4020
Db
3961 CTCTTCAATGATGCTGCTTCTTTCCAGATTCAGAAAGATTTCAATTTGATCAGGATTTGA 4020
QY
4021 ATGACCCCAATCCAGCTTTCAGCAGTGTGCTTGAACACTTAGGACAGAGAGAAAGAGC 4080
Db
4021 ATGACCCCAATCCAGCTTTCAGCAGTGTGCTTGAACACTTAGGACAGAGAGAAAGAGC 4080
QY
4081 CAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCCCATCCCATGCCGAACTTACA 4140
Db
4081 CAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCCCATCCCATGCCGAACTTACA 4140
QY
4141 CTGTGGCTGTTTATTTTCAAGGAGAGAAATAGCTGTGGGAAAGGACCAAGTATTCAGC 4200
Db
4141 CTGTGGCTGTTTATTTTCAAGGAGAGAAATAGCTGTGGGAAAGGACCAAGTATTCAGC 4200
QY
4201 AAGCGGAAATGGGAGCAGCAATGGATCGCTTGAAGAAATATAATTTTCCAGATGGCCC 4260
Db
4201 AAGCGGAAATGGGAGCAGCAATGGATCGCTTGAAGAAATATAATTTTCCAGATGGCCC 4260
QY
4261 ATCAGAAAGCGGTTTATCGAAACGGAAGTTACAGCAAGAGTTTAAAGAAATGAGGTGGGAA 4320
Db
4261 ATCAGAAAGCGGTTTATCGAAACGGAAGTTACAGCAAGAGTTTAAAGAAATGAGGTGGGAA 4320
QY
4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGAGGAGCAT 4380
Db
4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGAGGAGCAT 4380
QY
4381 GCAAGTGTGAGTATTTACTTGTCTGAGTAACTGTGACTGTTGCTATTTGAGACCTGACCT 4440
Db
4381 GCAAGTGTGAGTATTTACTTGTCTGAGTAACTGTGACTGTTGCTATTTGAGACCTGACCT 4440

[illegible][illegible]

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3725

```

Query Match	1.1%;	Score 54.4;	DB 6;	Length 3649;
Best Local Similarity	47.4%;	Pred. No. 0.00019;		
Matches 163;	Conservative 0;	Mismatches 181;	Indels 0;	Gaps 0;
3109	ATGACCCCAACTCCCTCGAGGATTAACCAACAATGAACGGTTGCGAATTCCTGGGGTGATCGTG	3168		
Db				
1787	ATGACTTTTAATATGATCGTTTGTAGAACACACAGCAACGGCTTAGAATTTTATGGTGATCGCG	1728		
Db				
3169	TTGTGGAATTCCTGACCGCGTCCATTTGTGTACTATTTGTTCCTTAGTCTGCGAAGAAGGAG	3228		
Db				
1727	TATTAGAAATGACGGTTTCACGCTATCTTTTGTGACAGACATCCCTCATTTACCAGAAGGTA	1668		
Db				
3229	GAATTAGCAACCTTATCGGACTCGCAATGTTTCAGATTCAGACACTTGGCCATGCTGACCAAGA	3288		
Db				
1667	ATTTGCAACAAGATGCGCGCAACAATTTGTTGTGCAACCTTCACTTTGTGATATTTTGCGAATA	1608		
Db				
3289	AACTTGAACTGGATCCATTTATGCTGTATGCTCACGGCGCTGACCTTTGTATGAGAAATCGG	3348		
Db				
1607	AGATTTAAATTAACGAACCTGATTTTATTAGGTAAAGGTGAAGAGACAGCAGGAGCGCAAGAA	1548		
Db				
3349	ACCTTCGCAATCAATGGCCAAATGTTTGTGAAGCGTTTAAATAGGAGCTGTTTACTTTCGGAGG	3408		
Db				
1547	CAAGACCTTCCCTTATTCAGATGCATTTGAAGCTTTTGTAGGTGCACCTGTATTAGATC	1488		
Db				
3409	GAAAGCCTGGAGGAAGCCAGCAGTATTATTGGACGGTTGCTCTTTT	3452		
Db				
1487	AAAGGTTTATGATTCAGTATGGACATTTGCTGAAAAGATCATCTTT	1444		
Db				

```

RESULT 5
US-11-121-086-18/c
; Sequence 18, Application US/1121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
; LENGTH: 175023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-18

```

	Query Match	1.1%;	Score 54.2;	DB 7;	Length 175023;
	Best Local Similarity	49.5%;	Pred. No. 0.0036;		
	Matches 140;	Conservative	0;	Mismatches 143;	Indels 0; Gaps 0;
QY	440	CTCTCCAGCCGCCCAATTTTCTCCCTCCACGACACAGACTTTGTATACCTTCCCTCCCTCCACCCCAT	499		
DB	123594	CTCCCTCTCCCTCCGCCCTCCCTCCCATCCCGCTCCCTCCCTCTCTCTCCCTCCCAT	123535		
QY	500	GCTCTCGGTGAGGCAAGGCGCTCTTCCCGCTGCGCCCAATCAGCGCGCGCTTTTCCCAACCA	559		
DB	123534	CGCCCTCGCTGCGCTCCCGCTTCCCTCCCGCTCCCTCCCTCCCTCCCTCCCTCCCTCT	123475		
QY	560	CCAGATGAGGCAACCCCTTCCCGAGTTCTCTCTTGTGTTTCTCCCATGCGCACCAATGCC	619		
DB	123474	CCCCCTCTCTCTCCCTCCCTTCTCTTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC	123415		
QY	620	TTGTCTCTAATAACCCCCCGGACCTTCTGGGCGGACCTTCTGGGACAGGCACTTTTCCCTCTCAT	679		

[illegible]

```

RESULT 7
US-11-121-086-2
; Sequence 2, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 191684
; TYPE: DNA

```

[illegible]

```

US-10-995-561-13498
Query Match      1.0%; Score 46.8; DB 6; Length 36360;
Best Local Similarity 46.8%; Pred. No. 0.14;
Matches 147; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 409 CTCGAAGTGCCTTCCACCACTTTCTCAAACTCTCAGGCCCGCAATTTCTCTCCCTCCAC 468
    |||||
Db 2209 CTCGCCCTCTCTCCCTCTCTCTCCCTCCCTCTCTCTCCCTCTCTCTCTCTCTCTCTCTCT 22150
    |||||

QY 469 GACCAGACTTTGTACCCCTTCCCGCCACCCATGCCTCGGTGAGCGAAGGCCCTCTTCCCC 528
    |||||
Db 22149 CCCTCCCTCTTTCCCGCCCTCTCCCTCTCTCTCCCTCTCTCTCCCGCCCTCTCTCCCTC 22090

QY 529 CDTGCCCAATCAGGCGCGCTTTTCCCAACACCAAGATGAGGCACCCCTTCCCAAGTTCCCTC 588
    |||||
Db 22089 CCCTTTCCTCTCCCGCTCTCCCTCTCTCTCCCGCTCTCTCTCTCTCTCTCTCTCTCTCT 22030

QY 589 CTGTGTTTCTCCCATGCCACCAAGTGCCTTGTCTTAATAACCCGCCAGTCTCTGGGG 648
    |||||
Db 22029 CTCCTCTCTCTCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21970

QY 649 CACCTCTCNGGCAAGGCACTTTCCCTTCATGATGCCCGCTCCCTCCATGCGCTCATCCCC 708
    |||||
Db 21969 CTCTGTCTCTCCCGCTCCCTCTCTCTCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21910

QY 709 GGCCCCCTCCAGTC 722
    |||||
Db 21909 TCCCTCCCTCCCTC 21896

RESULT 12
US-11-097-728-1
; Sequence 1, Application US/11097728
; Publication No. US20050260132A1
; GENERAL INFORMATION:
; APPLICANT: Erik Foehr
; APPLICANT: Sabine Muller
; APPLICANT: Daniel J. Chin
; APPLICANT: Mirella Gonzalez-Zulueta
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE ZETA
; FILE REFERENCE: AGYT-006CIP
; CURRENT APPLICATION NUMBER: US/11/097,728
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 10/652,981
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)...(7092)
US-11-097-728-1

Query Match      1.0%; Score 46.2; DB 7; Length 7941;
Best Local Similarity 49.4%; Pred. No. 0.069;
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 4522 GTTTTGTAGTCGATGGTCTTTATTAACAAGTATTAGATTTTCTTCTTCTTAAACGGAAA 4581
    |||||
Db 7693 GTAGTTCATTAGCTGGTCTTACTCTACCAGTTTCTCGACATTTGTTGTTACCTAAGT 7752
    |||||

QY 4582 CTTCACATTGTGTGAATGTCATTACTCTCTTTATTTTGTCTCTTTAAATAATAAATAATCA 4641
    |||||
Db 7753 CATTAACCTTTGTTTCAGCATGTAATTTTAACTTTTGTGGAAATAGAAATACCTTCATT 7812
    |||||

QY 4642 AGAAGCATATTTCTATGTGGAATAGATCCTGTTTTTCCATCTGTGTGCCAGATTGTGACCC 4701
    |||||
Db 7813 TGAAGAAGATTTTATGAGATAACACCTTACCAACCACTTGTTCAAAATCGTTTATATCCA 7872
    |||||

```

RESULT 10

US-11-121-086-53
; Sequence 53, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 53
; LENGTH: 176503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-53

Query Match 1.1%; Score 50.4; DB 7; Length 176503;
Best Local Similarity 47.2%; Pred. No. 0.043;
Matches 185; Conservative 0; Mismatches 206; Indels 1; Gaps 1;

Qy	408	CTTCAAAGTGCCCTTCACACACTTTCTCAAACTCTCAGGCCCCCAATTTTCTCCCTCCA	467
Db	112635	CGTTCCCGCGCCCTGCCCTCTTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	112694
Qy	468	CGACGACACTTTGTACCCCTTCCCACCACCATGCCTCGTCAGCGCAAGGCCCTCTTCCC	527
Db	112695	CTCTCCCTTATCTACACTCCCTCCCTTCCCTCCCAACCCCTCCCTCCCTCTCTC-ACCC	112753
Qy	528	CCCTGGCCAAATCAGCGCGCCTTTCCCAACACACAGATGAGGACACCCCTTCCCACTTCT	587
Db	112754	CTCCCTTTCCTCACCCCTCTCTTCCCTCCCTCTCTCTTCCCTCCCTCTCTCTTCCCTCA	112813
Qy	588	CCTTGTTTTCTCCGATGCGACGACCAATGCTTGCTATAAACCCGCGAGTCCCTGGG	647
Db	112814	CCCTCCCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCTCTC	112873
Qy	648	GCACTCTCTGACAAGGCATTTCCCTTCTCATGATGCCCTCCCTCCATGCTCATCCC	707
Db	112874	CCCTCCCTTCTCTTTCCCT	112933
Qy	708	CCGCCCCCTCCAGTCATGCCGACAGGTTAATTATCAGTACCCCTCCGGGCTATTCTCAC	767
Db	112934	CTGCCCCCTCCCTTCT	112993
Qy	768	CAGNACTTCCCACTCCCACTTTTAATAGTTT	799
Db	112994	CCCTCTCCCACT	113025

RESULT 11

US-10-995-561-13498/c
; Sequence 13498, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13498
; LENGTH: 36360
; TYPE: DNA
; ORGANISM: Homo sapiens

```
QY 4702 TAGACTTTCAATTGACAAAGTAAATAATGACTTTACTAGTAAATAAAAAAAAAAAAAA 4761
Db 7873 AGGAATTGCAAAATAAATAATAATGCAATTAATAAAAAAAAAAAAAAAAAAAAAA 7932
QY 4762 AAA 4764
Db 7933 AAA 7935

RESULT 13
US-11-097-728-5
; Sequence 5, Application US/11097728
; Publication No. US20050260132A1
; GENERAL INFORMATION:
; APPLICANT: Erik Foehr
; APPLICANT: Sabine Muller
; APPLICANT: Daniel J. Chin
; APPLICANT: Mirella Gonzalez-Zulueta
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN
; FILE REFERENCE: AGYT-006CIP
; CURRENT APPLICATION NUMBER: US/11/097,728
; PRIOR FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 10/652,981
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 8058
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (6229)...(6345)
; OTHER INFORMATION: Exon 23a splice variant of PTP-zeta SM2
; NAME/KEY: misc feature
; LOCATION: (7210)...(8058)
; OTHER INFORMATION: 3' untranslated region
US-11-097-728-5

Query Match 1.0%; Score 46.2; DB 7; Length 8058;
Best Local Similarity 49.4%; Pred. No. 0.07;
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 4522 GTTTTACTGATGCTCTTTATYACAAAGATTAGATTTTCTCTATTAAACGGAAA 4581
Db 7810 GTAGTTCATTAGCTGGTCTTACTCTACCAGTTTCTGACATGTATTGTGTACCCTAAGT 7869
QY 4582 CTTGACTTTGGTGAATGTGCATTACTCTCTTTATTGTCTTTTAAATAATAAAATTCA 4641
Db 7870 CATTAACTTTGTTTCAGCAGTGAATTTAACTTTGTGMAAATAGAAATACCTTCATT 7929
QY 4642 AGAAGCATATCTATGTGAATAGATCCCTGTTTTTCCATCTGTGTCCAGATTTGGACCC 4701
Db 7930 TGAAGAAGTTTTTATGAGAAATAACACCTTACCAACATTTGTTCAATGGTTTTTATCCA 7989
QY 4702 TAGACTTTCAATTGACAAAGTAAATAATGACTTTACTAGTAAATAAAAAAAAAAAAAA 4761
Db 7990 AGGAATTGCAAAATAAATAATAATTTGCCATTAAATAAAAAAAAAAAAAAAAAAAAAA 8049
QY 4762 AAA 4764
Db 8050 AAA 8052

RESULT 14
US-11-121-086-6/c
; Sequence 6, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
```

```
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 172543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-6

Query Match 1.0%; Score 45.8; DB 7; Length 172543;
Best Local Similarity 49.0%; Pred. No. 0.84;
Matches 151; Conservative 0; Mismatches 154; Indels 3; Gaps 1;

QY 425 CACCACCTTTCTCAAACTCTCCAGCCGCCCAATTTTCTCCCTCCAGCAGACGACTTTGTACC 484
Db 37135 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 37076
QY 485 CTTCCTCCCTCCATGCTCTCTCAGCGCAAGGCGCTCTTCCCTCCCTGCGCCCAATFCAAGCC 544
Db 37075 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 37016
QY 545 GCCTTTCCCAACACACAGATGAGGACCCCTTCCCTCCAGTCTCTCTCTCTCTCTCTCTCT 604
Db 37015 TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36956
QY 605 GCACACACCAA---TGCTTGTCTTAATAACCCCGCAGTCCCTGGGCGACCTCTCTGGACA 661
Db 36955 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36896
QY 662 AGGCACCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 721
Db 36895 CACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 721
QY 722 CATGCCGC 729
Db 36835 CTTCCTCAC 36828

RESULT 15
US-10-928-446A-1
; Sequence 1, Application US/10928446A
; Publication No. US2005027123A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
; TITLE OF INVENTION: VARIANTS OF NEED4L ASSOCIATED WITH HYPERTENSION AND
; TITLE OF INVENTION: VIRAL BUDDING
; FILE REFERENCE: 0274-5785.1US
; CURRENT APPLICATION NUMBER: US/10/928,446A
; CURRENT FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/359,741
; PRIOR FILING DATE: 2002-02-26
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1080000
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: allele
; LOCATION: (825234)
; OTHER INFORMATION: the 'n' at position 825234 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (825270)
; OTHER INFORMATION: the 'n' at position 825270 may be 'c' or 'g'
; FEATURE:
; NAME/KEY: allele
```



```

; LOCATION: (825401)
; OTHER INFORMATION: the 'n' at position 825401 may be 'c' or 'a'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (825428)
; OTHER INFORMATION: the 'n' at position 825428 may be 'g' or 'a'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (825473)
; OTHER INFORMATION: the 'n' at position 825473 may be 'g' or 'a'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (825765)
; OTHER INFORMATION: a "c" may be deleted at this position
; FEATURE:
; NAME/KEY: allele
; LOCATION: (825828)
; OTHER INFORMATION: the 'n' at position 825828 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (826041)
; OTHER INFORMATION: the 'n' at position 826041 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (826546)
; OTHER INFORMATION: the 'n' at position 826546 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (826654)
; OTHER INFORMATION: the 'n' at position 826654 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (826826)
; OTHER INFORMATION: the 'n' at position 826826 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (826863)
; OTHER INFORMATION: the 'n' at position 826863 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (827008)
; OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (843055)
; OTHER INFORMATION: the 'n' at position 843055 may be 'g' or 'a'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (843118)
; OTHER INFORMATION: the 'n' at position 843118 may be 't' or 'c'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (871027)
; OTHER INFORMATION: the 'n' at position 871027 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (871140)
; OTHER INFORMATION: the 'n' at position 871140 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (871168)
; OTHER INFORMATION: the 'n' at position 871168 may be 'c' or 'a'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (872678)
; OTHER INFORMATION: the 'n' at position 872678 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (872742)
; OTHER INFORMATION: the 'n' at position 872742 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (925859)

```

```

; OTHER INFORMATION: the 'n' at position 925859 may be 'c' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (993220)
; OTHER INFORMATION: the 'n' at position 993220 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (993254)
; OTHER INFORMATION: the 'n' at position 993254 may be 'g' or 'a'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1006462)
; OTHER INFORMATION: the 'n' at position 1006462 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1007820)
; OTHER INFORMATION: "ttct" may be deleted at this position
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1018038)
; OTHER INFORMATION: the 'n' at position 1018038 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1018704)
; OTHER INFORMATION: the 'n' at position 1018704 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1018718)
; OTHER INFORMATION: "gtt" may be deleted at this position
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1026786)
; OTHER INFORMATION: the 'n' at position 1026786 may be 'c' or 'a'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1047134)
; OTHER INFORMATION: the 'n' at position 1047134 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1047159)
; OTHER INFORMATION: the 'n' at position 1047159 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1047378)
; OTHER INFORMATION: the 'n' at position 1047378 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1047739)
; OTHER INFORMATION: the 'n' at position 1047739 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1050133)
; OTHER INFORMATION: "ttaa" may be deleted at this position
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1050539)
; OTHER INFORMATION: the 'n' at position 1050539 may be 'c' or 't'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1062808)
; OTHER INFORMATION: the 'n' at position 1062808 may be 'c' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1066392)
; OTHER INFORMATION: the 'n' at position 1066392 may be 'a' or 'g'
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1073711)
; OTHER INFORMATION: the 'n' at position 1073711 may be 'c' or 't'
; FEATURE:
; US-10-928-446A-1

```

```

Query Match      1.0%; Score 45.6; DB 6; Length 1080000;
Best Local Similarity 49.0%; Pred. No. 3.6;

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 14:23:08 ; Search time 23208 Seconds
(without alignments)
11668.485 Million cell updates/sec

Title: US-10-774-974-1
Perfect score: 4764
Sequence: 1 ctgtcttggtacctgcggta.....aaaaaaaaaaaaaaaaaaaaa 4764

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5863141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4764	100.0	4764	6 AR540925	Sequence
2	4760.8	99.9	4764	6 CS033685	Sequence
3	4760.8	99.9	4764	6 CS042637	Sequence
4	4760.8	99.9	4764	8 AF189011	Homo sapi
5	4343.4	91.2	4650	8 BX647724	Homo sapi
6	3600	75.6	4069	8 BC054003	Homo sapi
7	3379.6	70.9	4479	9 BC088999	Mus muscu
8	3042.2	63.9	3838	9 BC060265	Mus muscu
9	2582	54.2	2598	6 BD155916	Primer fo
10	2582	54.2	2598	6 AX876043	Sequence
11	2582	54.2	2598	8 AK001121	Homo sapi
12	2564	53.8	4516	5 AJ720721	Gallus ga
13	2525.2	53.0	2607	8 BC024261	Homo sapi
14	2442	51.3	2473	8 BC041162	Homo sapi
15	2425.6	50.9	2971	9 BC057687	Mus muscu
16	2425.2	50.9	3066	9 AF533013	Mus muscu
17	2272	47.7	2292	8 AF116910	Homo sapi
18	2032.4	42.7	2488	9 BC055696	Mus muscu

19	1829.8	38.4	2348	9 BC050057	Mus muscu
20	1827.2	38.4	1836	6 CQ720739	Sequence
21	1744	36.6	2229	9 BC040801	Mus muscu
22	1619.6	34.0	1626	6 AX700231	Sequence
23	1607.8	33.7	2094	6 CQ720503	Sequence
24	1496	31.4	1518	8 AK026563	Homo sapi
25	1440.8	30.2	1743	9 AF440766	Mus muscu
26	1376	26.8	11720	8 AB209501	Homo sapi
27	897.2	18.8	11443	8 HSA242976	Homo sapi
28	832.2	17.5	124028	8 AC022417	Homo sapi
29	818.8	17.2	897	6 BD146328	Primer fo
30	818.8	17.2	897	6 AX866266	Sequence
31	770.2	16.2	3825	2 AY051775	Drosophil
32	770.2	16.2	4275	6 CQ574650	Sequence
33	770.2	16.2	4286	2 AF116572	Drosophil
34	770.2	16.2	4311	2 BT023503	Drosophil
35	741.2	15.6	995	9 AY373464	Rattus no
36	725.4	15.2	1003	6 AX053322	Sequence
37	724	15.2	803	6 AX053392	Sequence
38	703.8	14.8	6397	6 CQ574649	Sequence
39	703.8	14.8	85651	14 AC005463	Drosophil
40	703.8	14.8	86889	2 AC005448	Drosophil
41	703.8	14.8	139410	2 AC007084	Drosophil
42	703.8	14.8	254177	2 AB003839	Drosophil
43	693.4	14.6	37983	14 AC020394	Drosophil
44	683.4	14.3	833	6 AX053394	Sequence
45	602.4	12.6	763	6 AX053323	Sequence

ALIGNMENTS

RESULT 1
AR540925
LOCUS AR540925 4764 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6737512.
ACCESSION AR540925
VERSION AR540925.1 GI:53932478
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4764)
AUTHORS Wu,H. and Crooke,S.T.
TITLE Human RNase III and compositions and uses thereof
JOURNAL Patent: US 6737512-A 1 18-MAY-2004;
ISIS Pharmaceuticals, Inc.; Carlebad, CA
FEATURES
source location/Qualifiers
1..4764
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match	100.0%;	Score 4764;	DB 6;	Length 4764;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 4764;	Conservative 0;	0;	0;	Gaps 0;
Qy	1	CTGCTCTGTACTCGCGGTAGTAGCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG	60	
Db	1	CTGCTCTGTACTCGCGGTAGTAGCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG	60	
Qy	61	AGCCTTTTATAGTTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	120	
Db	61	AGCCTTTTATAGTTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	120	
Qy	121	CCCATATCATCAAGAGCTGATATCTAGTGAAGATTAGAGTTAGCGTGTGCATCTTCACTA	180	
Db	121	CCCATATCATCAAGAGCTGATATCTAGTGAAGATTAGAGTTAGCGTGTGCATCTTCACTA	180	
Qy	181	TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGCATATCCAGCGGA	240	
Db	181	TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGCATATCCAGCGGA	240	

QY	241	ACATCATGATGAGGGAAACA	CATGTACAGAAATGTGTTCCACCCCGGAGCAGAGGGCGTC	300				
DB	241	ACATCATGATGAGGGAAACA	CATGTACAGAAATGTGTTCCACCCCGGAGCAGAGGGCGTC	300				
QY	301	CCCGAGGACGAGGAGGACA	TGAGGCCAGACCCTCAGACCATCTTTTAGGCCCCCAAAATC	360				
DB	301	CCCGAGGACGAGGAGGACA	TGAGGCCAGACCCTCAGACCATCTTTTAGGCCCCCAAAATC	360				
QY	361	TGAGGCTGCTTCA	CCCTCAGCAGCCTCTGTGCAATATCAATATGAACCTTCCAAGTGCCCC	420				
DB	361	TGAGGCTGCTTCA	CCCTCAGCAGCCTCTGTGCAATATCAATATGAACCTTCCAAGTGCCCC	420				
QY	421	CTTCCACA	CACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCAGCAGCAGACTTTG	480				
DB	421	CTTCCACA	CACTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCAGCAGCAGACTTTG	480				
QY	481	TACCTTCCCCCACCACCT	CGTCCGTGAGCGAGGCCCTCTTCCCCCTGCCCAATCA	540				
DB	481	TACCTTCCCCCACCACCT	CGTCCGTGAGCGAGGCCCTCTTCCCCCTGCCCAATCA	540				
QY	541	GGCCGCTTTCCCAACCA	CACAGATGAGGCACCCCTTCCCAGTTCTCTCTGTTGTTTCCCTC	600				
DB	541	GGCCGCTTTCCCAACCA	CACAGATGAGGCACCCCTTCCCAGTTCTCTCTGTTGTTTCCCTC	600				
QY	601	CCATGCCACCA	CCAAATGCTTGTCTAATAACCCCCAGTCCCTCTGGGCACCTCTGGAC	660				
DB	601	CCATGCCACCA	CCAAATGCTTGTCTAATAACCCCCAGTCCCTCTGGGCACCTCTGGAC	660				
QY	661	AAGGCAC	TTTCCCCTCATGATGCCCTCCCTCCATGCCTCATCCCGCCGCCCTCCAG	720				
DB	661	AAGGCAC	TTTCCCCTCATGATGCCCTCCCTCCATGCCTCATCCCGCCGCCCTCCAG	720				
QY	721	TCATCCCGCAGCAGGT	TAATATCAGTACCTTCGGGCTATTCTCACACAACCTTCTCCAC	780				
DB	721	TCATCCCGCAGCAGGT	TAATATCAGTACCTTCGGGCTATTCTCACACAACCTTCTCCAC	780				
QY	781	CTCCAGTTT	TAATAGTTTCAGAAACAACCTTAGTTCTTCTGCCAGTGTCTAATAACA	840				
DB	781	CTCCAGTTT	TAATAGTTTTCAGAAACAACCTTAGTTCTTCTGCCAGTGTCTAATAACA	840				
QY	841	GCAGTAGTCTCAT	TTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAAGTGAGA	900				
DB	841	GCAGTAGTCTCAT	TTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAAGTGAGA	900				
QY	901	GAAAGTCCCCA	AGAGGCTGAAACA	CTATGATGACCA	CAGGCACCGGAGAC	CACAGTCA	TG	960
DB	901	GAAAGTCCCCA	AGAGGCTGAAACA	CTATGATGACCA	CAGGCACCGGAGAC	CACAGTCA	TG	960
QY	961	GGCAGGTGAGAGG	CATCGTCTCTGGATCGCGCGGAGCGAGGCCGAGTCCCGACAGGA	1020				
DB	961	GGCAGGTGAGAGG	CATCGTCTCTGGATCGCGCGGAGCGAGGCCGAGTCCCGACAGGA	1020				
QY	1021	GAAACA	GACAGCGGTACAGATCTGATTAATGACCGAGGGAGAACACCATCTCGCCACC	1080				
DB	1021	GAAACA	GACAGCGGTACAGATCTGATTAATGACCGAGGGAGAACACCATCTCGCCACC	1080				
QY	1081	GCAGTCTACGAA	CGGAGCAGAGCCAGAACCGGAGAGACACAGGCATCTCGACAACCCGAA	1140				
DB	1081	GCAGTCTACGAA	CGGAGCAGAGCCAGAACCGGAGAGACACAGGCATCTCGACAACCCGAA	1140				
QY	1141	GATCACC	ATCTCTGGAAGGTCCTCA	AAAAAGAGTAT	AAGAGAT	TGGAAGGAT	TACG	1200
DB	1141	GATCACC	ATCTCTGGAAGGTCCTCA	AAAAAGAGTAT	AAGAGAT	TGGAAGGAT	TACG	1200
QY	1201	GTTTATCGGTTG	TCTGAA	CCTGCTGATGCA	CACAGAT	TACCTGGGGAGAT	TATTA	1260
DB	1201	GTTTATCGGTTG	TCTGAA	CCTGCTGATGCA	CACAGAT	TACCTGGGGAGAT	TATTA	1260
QY	1261	AAAA	TACAGATTCTTTGGGCCCC	CACCCCTGGAGATTGTGAATCAT	CGCTCCCAAGTAGGG	1320		
DB	1261	AAAA	TACAGATTCTTTGGGCCCC	CACCCCTGGAGATTGTGTGAATCAT	CGCTCCCAAGTAGGG	1320		
QY	1321	AGAAGA	GAGAGCTCTGTTGGGAGGAGAAAAAG	ACCGTTGGAGTGACA	ACACAGAGTTCTG	1380		

1321	AGAAAGAGAGCTCGTTGGAGAGAAAGACCGTTGGAGTGCAACCGAGATTCGT	1381
1381	GCAAGACAAGAACTATACCTCAATCAAGSAAAAAGAGCCGAGGAGACCATGCTGACA	1440
1381	GCAAGACAAGAACTATACCTCAATCAAGSAAAAAGAGCCGAGGAGACCATGCTGACA	1440
1441	AGAAAGAGAGAGAGAGAGAACTTCCTTAAGCTGTGTGGATTGCGATGCACTCATTCAG	1500
1441	AGAAAGAGAGAGAGAGAGAACTTCCTTAAGCTGTGTGGATTGCGATGCACTCATTCAG	1500
1501	AAAACTACTACTCCAGTGAACCCATGATCAGGTGGAGATTCTACGTGGTTGGAAAGA	1560
1501	AAAACTACTACTCCAGTGAACCCATGATCAGGTGGAGATTCTACGTGGTTGGAAAGA	1560
1561	GTAGGCTTCGTGACTTATATGACAAAATTGAGGAGAGATTGGGGAGCAGGCAAGAAAGG	1620
1561	GTAGGCTTCGTGACTTATATGACAAAATTGAGGAGAGATTGGGGAGCAGGCAAGAAAGG	1620
1621	CAAAAGCTGCTCGGCTCCGTTGGAGAACTCTCAAGACGAAAGCTCGATGAAGATTAGAGA	1680
1621	CAAAAGCTGCTCGGCTCCGTTGGAGAACTCTCAAGACGAAAGCTCGATGAAGATTAGAGA	1680
1681	GTTCAGATGAATCCGAGTGTGAGTGTGATGAGACAGACCTGTGTTAGCAGCTCAGACT	1740
1681	GTTCAGATGAATCCGAGTGTGAGTGTGATGAGACAGACCTGTGTTAGCAGCTCAGACT	1740
1741	CTGAAGTTTGTGAGTTATTCAGAAATCAAAACGCAAAAAGGCCACCCCTGACCGACTTC	1800
1741	CTGAAGTTTGTGAGTTATTCAGAAATCAAAACGCAAAAAGGCCACCCCTGACCGACTTC	1800
1801	ATGATGAATCTTTGGTGAACGATCCAGGCGAGATGAATGATGGAACACTCTGCAAAATGCA	1860
1801	ATGATGAATCTTTGGTGAACGATCCAGGCGAGATGAATGATGGAACACTCTGCAAAATGCA	1860
1861	GCGCAAGGCAAGCGCACAGGAATTAGGCACAGCACTTATCCTCGAGAAAGGCGCATCA	1920
1861	GCGCAAGGCAAGCGCACAGGAATTAGGCACAGCACTTATCCTCGAGAAAGGCGCATCA	1920
1921	AGCCCTGTGCTCCTATGACCAAACTGCTGGCAGACTTTTCCACTACCGGATCAAGTCT	1980
1921	AGCCCTGTGCTCCTATGACCAAACTGCTGGCAGACTTTTCCACTACCGGATCAAGTCT	1980
1981	CCCGGCTACGAACTTTTAACTGACAGGCGAACTGTGTATAGAAATCGATGATCAAGT	2040
1981	CCCGGCTACGAACTTTTAACTGACAGGCGAACTGTGTATAGAAATCGATGATCAAGT	2040
2041	ATATCTTTGAGGATTTTCTATGTTGCAATGCCCTCGACCAATATTCACATGTGTA	2100
2041	ATATCTTTGAGGATTTTCTATGTTGCAATGCCCTCGACCAATATTCACATGTGTA	2100
2101	AGTAATTAGATTCAACATAGACTACAGATTCAITTTCAATTGAAGAGATGATGCGGAGA	2160
2101	AGTAATTAGATTCAACATAGACTACAGATTCAITTTCAATTGAAGAGATGATGCGGAGA	2160
2161	ATTTTGTGTGAAGGGCTTGAACTCTTTTCACTGTTCCTATTCAGAGAAATTTTGGAA	2220
2161	ATTTTGTGTGAAGGGCTTGAACTCTTTTCACTGTTCCTATTCAGAGAAATTTTGGAA	2220
2221	TATATGACTGGAACTTTAAAGGTCCTTTGTTTGAAGCAGCCCTCCCTGCTGCCCAAGT	2280
2221	TATATGACTGGAACTTTAAAGGTCCTTTGTTTGAAGCAGCCCTCCCTGCTGCCCAAGT	2280
2281	TTCAITTCATGCGAGTTTCTGAAGATTCTTCCAGATGGAGSAAAGGAAGTGTGTC	2340
2281	TTCAITTCATGCGAGTTTCTGAAGATTCTTCCAGATGGAGSAAAGGAAGTGTGTC	2340
2341	TGCAACAGATTCCTCTGATCTTTAAGGTGCAAGCCCTGTGTCTGAGGAGAGA	2400
2341	TGCAACAGATTCCTCTGATCTTTAAGGTGCAAGCCCTGTGTCTGAGGAGAGA	2400
2401	TTGCCAATATGCTTCAGTGGGAGGAGCTGAGTGGCAGAAATATGACAGAAATCAAG	2460

Db	2401	TTGCCAATATGCTTCAGTGGGAGCGTGGAGTGGCAGAAATATGCAAGAATGCAAG	2460
Qy	2461	GCATGATTTGTACCAACCTTGGGACGAAACCAAGCTCTGTCGTATCGATCAACTGGATC	2520
Db	2461	GCATGATTTGTACCAACCTTGGGACGAAACCAAGCTCTGTCGTATCGATCAACTGGATC	2520
Qy	2521	GTGAACAGTTCAACCCCGATGTGATTTCTCCGATTAATCGTCCACTTTGGGATACGCC	2580
Db	2521	GTGAACAGTTCAACCCCGATGTGATTTCTTTCCGATTAATCGTCCACTTTGGGATACGCC	2580
Qy	2581	CTGCACAGTTGAGTTATGACGAGAGACCCACAGTACCAGAACTGTGGAGAGTTATGTGA	2640
Db	2581	CTGCACAGTTGAGTTATGACGAGAGACCCACAGTACCAGAACTGTGGAGAGTTATGTGA	2640
Qy	2641	AACCTCGCCACTCTAGCAAAATAGTCCCAAAGTCAAACTGACAAACAGAGCTGG	2700
Db	2641	AACCTCGCCACTCTAGCAAAATAGTCCCAAAGTCAAACTGACAAACAGAGCTGG	2700
Qy	2701	CACAGAGGAGGAGCCCTCCAAAATAACGGCAGAGAGATACAATGACAGAGAGTAA	2760
Db	2701	CACAGAGGAGGAGGAGCCCTCCAAAATAACGGCAGAGAGATACAATGACAGAGAGTAA	2760
Qy	2761	CGGTGAGCTAAGTAGCCAAAGGATCTGGAAAACTGGCATCCGTTCTGATGCTGTGAGC	2820
Db	2761	CGGTGAGCTAAGTAGCCAAAGGATCTGGAAAACTGGCATCCGTTCTGATGCTGTGAGC	2820
Qy	2821	ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCAAT	2880
Db	2821	ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCAAT	2880
Qy	2881	TGGACAAAGTTGATAGATATACCTTCCAGAGTGGTGTGTCGTGACGCTGGCCATGACTC	2940
Db	2881	TGGACAAAGTTGATAGATATACCTTCCAGAGTGGTGTGTCGTGACGCTGGCCATGACTC	2940
Qy	2941	ATCCAAAGTCAATTTAAATTTTGGAAATGAATCTGATGATGCGAGAAATTCATATCTA	3000
Db	2941	ATCCAAAGTCAATTTAAATTTTGGAAATGAATCTGATGATGCGAGAAATTCATATCTA	3000
Qy	3001	ACTGTGGAATTCGGCAGCCCAATAACGAGACAGAAAGTTCAATCAATGCAATCGGA	3060
Db	3001	ACTGTGGAATTCGGCAGCCCAATAACGAGACAGAAAGTTCAATCAATGCAATCGGA	3060
Qy	3061	AGAAAGGATTAACCTGTTGATAAATATCATGTGACGCTTGGCCAAAGTACCCAACTC	3120
Db	3061	AGAAAGGATTAACCTGTTGATAAATATCATGTGACGCTTGGCCAAAGTACCCAACTC	3120
Qy	3121	CCTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTGTAATTC	3180
Db	3121	CCTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTGTAATTC	3180
Qy	3181	TGACCAAGCTTCAATTTGTAATTTGTTCTAGTCTGGAAGAGGAGGATTAGCAACCT	3240
Db	3181	TGACCAAGCTTCAATTTGTAATTTGTTCTAGTCTGGAAGAGGAGGATTAGCAACCT	3240
Qy	3241	ATCGGATCGCATTTGTAATCAGCACTTCCCATGCTAGCAAGAACTTGAATCGG	3300
Db	3241	ATCGGATCGCATTTGTAATCAGCACTTCCCATGCTAGCAAGAACTTGAATCGG	3300
Qy	3301	ATCCATTTATGCTGATGCTCACGGGCTGACCTTTGTAGAGAACTCGACCTTCGACATG	3360
Db	3301	ATCCATTTATGCTGATGCTCACGGGCTGACCTTTGTAGAGAACTCGACCTTCGACATG	3360
Qy	3361	CAATGGCCAAATGTTTTGAAGCTTAATAGAGCTGTTTCTTTGGAGGAGGAGCTGGAGG	3420
Db	3361	CAATGGCCAAATGTTTTGAAGCTTAATAGAGCTGTTTCTTTGGAGGAGGAGCTGGAGG	3420
Qy	3421	AAGCCAAAGGATTTATGAGCGTCTGCTCTTTAATGATCGGACCTTGGGAACTGCGC	3480
Db	3421	AAGCCAAAGGATTTATGAGCGTCTGCTCTTTAATGATCGGACCTTGGGAACTGCGC	3480
Qy	3481	TCAATTTATCTCTCCACCACTCAACTACAGAGCCAAATATCTGATCGCAACTTTATG	3540
Db	3481	TCAATTTATCTCTCCACCACTCAACTACAGAGCCAAATATCTGATCGCAACTTTATG	3540

Qy	3541	AAACTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATTTGGAGTAATTTTAA	3600
Db	3541	AAACTTCTCCAGTTCTACAAAACCTTACTGAGTTTGAAGAGCAATTTGGAGTAATTTTAA	3600
Qy	3601	CTCATGTTTCGACTTCTTGGCAAGGGCAATTCACATTTGAGAACTGTGGGATTTAAACCATCTGA	3660
Db	3601	CTCATGTTTCGACTTCTTGGCAAGGGCAATTCACATTTGAGAACTGTGGGATTTAAACCATCTGA	3660
Qy	3661	CCCTAGGCCACAATTCAGAGAAATGGAATTCCTAGGTGACTTCCATAATATGCAACTGGTACCCA	3720
Db	3661	CCCTAGGCCACAATTCAGAGAAATGGAATTCCTAGGTGACTTCCATAATATGCAACTGGTACCCA	3720
Qy	3721	CAGAGTACTTATTCATTTCAATTTCCAGATCCTATGAAGGACACTTAACTTTGTTGGGAA	3780
Db	3721	CAGAGTACTTATTCATTTCAATTTCCAGATCCTATGAAGGACACTTAACTTTGTTGGGAA	3780
Qy	3781	GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGTACGCGAGGAGCTGGGCATGCAGAGT	3840
Db	3781	GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGTACGCGAGGAGCTGGGCATGCAGAGT	3840
Qy	3841	AGCCCAATACCAACGACCAAGAGGCTGTGGCGCTTCGCAACAAGACCTTGGCGG	3900
Db	3841	AGCCCAATACCAACGACCAAGAGGCTGTGGCGCTTCGCAACAAGACCTTGGCGG	3900
Qy	3901	ACCTTTTGGAAATTCATTTATGAGCGCTGTACATGATAGGATTTGGAAATATGTTTCATA	3960
Db	3901	ACCTTTTGGAAATTCATTTATGAGCGCTGTACATGATAGGATTTGGAAATATGTTTCATA	3960
Qy	3961	CTTTCAATGAATGTCCTCTTTCCACGATTTGAAGAAATTCATTTTGAATCAGGATTTGA	4020
Db	3961	CTTTCAATGAATGTCCTCTTTCCACGATTTGAAGAAATTCATTTTGAATCAGGATTTGA	4020
Qy	4021	ATGACCCCAAAATCCAGCTTCAGCAGTGTGCTTGAACAATTTAGGAGTAAAGAGCA	4080
Db	4021	ATGACCCCAAAATCCAGCTTCAGCAGTGTGCTTGAACAATTTAGGAGTAAAGAGCA	4080
Qy	4081	CAGACATTTCTCTGTACAGACTCTTCAGACAGTGGGCGCCATCCCATGCCGCACTTACA	4140
Db	4081	CAGACATTTCTCTGTACAGACTCTTCAGACAGTGGGCGCCATCCCATGCCGCACTTACA	4140
Qy	4141	CTGTGCTGTTTATTTTCAAGGGAGAAAGTATAGGCTGTGGGAAAGGACCAAGTATTCAGC	4200
Db	4141	CTGTGCTGTTTATTTTCAAGGGAGAAAGTATAGGCTGTGGGAAAGGACCAAGTATTCAGC	4200
Qy	4201	AAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAATAATAATTTTCCCAGATGSCCC	4260
Db	4201	AAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAATAATAATTTTCCCAGATGSCCC	4260
Qy	4261	ATCAGAAAGCGGTTTCATCGAAACGGAAGTACAGACAGTGTAAAGAAATGAGGTGGGAAA	4320
Db	4261	ATCAGAAAGCGGTTTCATCGAAACGGAAGTACAGACAGTGTAAAGAAATGAGGTGGGAAA	4320
Qy	4321	GAGAGCATCAAGAGAGAGAGCCAGATGAGATGCAAGACATCAAGAAATTAAGAGGGGCAT	4380
Db	4321	GAGAGCATCAAGAGAGAGAGCCAGATGAGATGCAAGACATCAAGAAATTAAGAGGGGCAT	4380
Qy	4381	GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGTGCTGTTGTCTATTTGAGACCTAGCCT	4440
Db	4381	GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGTGCTGTTGTCTATTTGAGACCTAGCCT	4440
Qy	4441	AGTTTTCTCGCAGACAATGAACGAAAGTGTCTATTTAGTGAATTAATAACAGAGTCAATTCG	4500
Db	4441	AGTTTTCTCGCAGACAATGAACGAAAGTGTCTATTTAGTGAATTAATAACAGAGTCAATTCG	4500
Qy	4501	CTATTTCTTAAATGATCTGTTTTAGTGTGATGCTCTTTATTAACAAGATTTAGATTT	4560
Db	4501	CTATTTCTTAAATGATCTGTTTTAGTGTGATGCTCTTTATTAACAAGATTTAGATTT	4560
Qy	4561	TTTTCTTCTTAAACGGAACCTTGAATTTGGTGAATTTGCACTTCTTTTATTTTG	4620
Db	4561	TTTTCTTCTTAAACGGAACCTTGAATTTGGTGAATTTGCACTTCTTTTATTTTG	4620

QY 4621 CTCCTTAAATAATAAAATTCAGAAAGCATATCTATGTGGAATAGATCCTGTTTTCAT 4680
Db |||||||
QY 4621 CTCCTTAAATAATAAAATTCAGAAAGCATATCTATGTGGAATAGATCCTGTTTTCAT 4680
Db |||||||
QY 4681 CTGTGTCCAGATTTGACCCCTAGACTTTCAATTGACAAAGTAAAAATTTGACTTTACTAG 4740
Db |||||||
QY 4681 CTGTGTCCAGATTTGACCCCTAGACTTTCAATTGACAAAGTAAAAATTTGACTTTACTAG 4740
Db |||||||
QY 4741 TAAAAAATAAAAAAAAAAAAAA 4764
Db |||||||
QY 4741 TAAAAAATAAAAAAAAAAAAAA 4764
Db |||||||

RESULT 2
CS033685 LOCUS CS033685 4764 bp DNA linear PAT 10-MAR-2005
DEFINITION Sequence 3191 from Patent WO2005016962.
ACCESSION CS033685
VERSION CS033685.1 GI:60732625
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE Abbas, A., Clark, H., Ouyang, W., Williams, M. P., Wood, W. I. and Wu, T. D.
AUTHORS Compositions and methods for the treatment of immune related
TITLE diseases
JOURNAL Patent: WO 2005016962-A 3191 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
1..4764
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.9%; Score 4760.8; DB 6; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGTCTTGTTGCTACCTCGGTAGTAGCTGCTGCTCTGACGCGCATCTCGGGCCCGAG 60
Db |||||||
QY 1 CTGTCTTGTTGCTACCTCGGTAGTAGCTGCTGCTCTGACGCGCATCTCGGGCCCGAG 60
Db |||||||
QY 61 AGCCTTTTATAGTTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
Db |||||||
QY 61 AGCCTTTTATAGTTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
Db |||||||
QY 121 CCATATCATCAAGGAGCTGATATCTAGTGGAGAGTTAGAGCTGTGCATCTTCACTA 180
Db |||||||
QY 121 CCATATCATCAAGGAGCTGATAATCTAGTGGAGAGTTAGAGCTGTGCATCTTCACTA 180
Db |||||||
QY 181 TGAATAGAGCAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGCGGA 240
Db |||||||
QY 181 TGAATAGAGCAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGCGGA 240
Db |||||||
QY 241 ACATCATGATGACGGGAACACATGTGCACAGATGTCTGTCCACCGGACGAGGCGTC 300
Db |||||||
QY 241 ACATCATGATGACGGGAACACATGTGCACAGATGTCTGTCCACCGGACGAGGCGTC 300
Db |||||||
QY 301 CCGAGGACGAGGAGACATGAGACACCTTCAGACCATCTTTAGGCCCCCAAAATC 360
Db |||||||
QY 301 CCGAGGACGAGGAGACATGAGACACCTTCAGACCATCTTTAGGCCCCCAAAATC 360
Db |||||||
QY 361 TGAGGCTGCTTCAACCTTCAGAGCCCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
Db |||||||
QY 361 TGAGGCTGCTTCAACCTTCAGAGCCCTCTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
Db |||||||
QY 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACGACTTTG 480
Db |||||||
QY 421 CTTTCCACCACTTTCTCAAACTCTCCAGCCCCCAATTTCTCCCTCCACGACGACTTTG 480
Db |||||||

QY 481 TACCCTTCCCCCACCATGCTCCGTAGCGCAAGGCCCTCTTCCCCCTGCCAATCA 540
Db |||||||
QY 481 TACCCTTCCCCCACCATGCTCCGTAGCGCAAGGCCCTCTTCCCCCTGCCAATCA 540
Db |||||||
QY 541 GGCCGCTTTTCCCCAACCAACAGATGAGGCAACCCCTTCCAGTTCTCTGTTTCTTC 600
Db |||||||
QY 541 GGCCGCTTTTCCCCAACCAACAGATGAGGCAACCCCTTCCAGTTCTCTGTTTCTTC 600
Db |||||||
QY 601 CCATGCGACCAATGCTTGTCTTAATAACCCGCCAGTCCCTGGGGCACCTCTCTGAC 660
Db |||||||
QY 601 CCATGCGACCAATGCTTGTCTTAATAACCCGCCAGTCCCTGGGGCACCTCTCTGAC 660
Db |||||||
QY 661 AAGGCACTTTTCCCTTCATGATGCCCTCCCTCCATGCTCATCCCCCGCCCTCCAG 720
Db |||||||
QY 661 AAGGCACTTTTCCCTTCATGATGCCCTCCCTCCATGCTCATCCCCCGCCCTCCAG 720
Db |||||||
QY 721 TCATGCCGAGCAGGTTAATTATCAGTACCTTCCGGGCTATTCTCAACCACTTCCAC 780
Db |||||||
QY 721 TCATGCCGAGCAGGTTAATTATCAGTACCTTCCGGGCTATTCTCAACCACTTCCAC 780
Db |||||||
QY 781 CTCACGATTTTAAATAGTTTCCAGAACCCCTAGTTCTTCTGCCCAGTGCTTAATA 840
Db |||||||
QY 781 CTCACGATTTTAAATAGTTTCCAGAACCCCTAGTTCTTCTGCCCAGTGCTTAATA 840
Db |||||||
QY 841 GCAAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAAGTGA 900
Db |||||||
QY 841 GCAAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAAGTGA 900
Db |||||||
QY 901 GAAGGTCCCAGAAAGGCTGAAACATATGATGACACAGGACCCGAGACCAAGTATG 960
Db |||||||
QY 901 GAAGGTCCCAGAAAGGCTGAAACATATGATGACACAGGACCCGAGACCAAGTATG 960
Db |||||||
QY 961 GCGAGGTGAGAGGATCGGTCCCTGATCGCGGGAGGCGGCGAGTCCCGACGGA 1020
Db |||||||
QY 961 GCGAGGTGAGAGGATCGGTCCCTGATCGCGGGAGGCGGCGAGTCCCGACGGA 1020
Db |||||||
QY 1021 GAAGCAAGACAGCCGCTACAGATCTGATTTATGACCGAGGAGAACACCATCTCGCCACC 1080
Db |||||||
QY 1021 GAAGCAAGACAGCCGCTACAGATCTGATTTATGACCGAGGAGAACACCATCTCGCCACC 1080
Db |||||||
QY 1081 GCAGCTACGAAACGAGAGCGAGAACCGGAGAGACACAGGATCGAGACCAACCGAA 1140
Db |||||||
QY 1081 GCAGCTACGAAACGAGAGCGAGAACCGGAGAGACACAGGATCGAGACCAACCGAA 1140
Db |||||||
QY 1141 GATCACCATCTCTGAAAGTCTCTACAAAGAGTATGAGATCTGAAAGGATGAG 1200
Db |||||||
QY 1141 GATCACCATCTCTGAAAGTCTCTACAAAGAGTATGAGATCTGAAAGGATGAG 1200
Db |||||||
QY 1201 GTTTATCGGTCTTCTGAACTCTGATGACACAGAAATTTACCTGGGGAGATTATTA 1260
Db |||||||
QY 1201 GTTTATCGGTCTTCTGAACTCTGATGACACAGAAATTTACCTGGGGAGATTATTA 1260
Db |||||||
QY 1261 AAAATACAGATCTTTGGGCCCCACCCCTGGAGATTTGAAATCATCGCTCCCAAGTAGG 1320
Db |||||||
QY 1261 AAAATACAGATCTTTGGGCCCCACCCCTGGAGATTTGAAATCATCGCTCCCAAGTAGG 1320
Db |||||||
QY 1321 AGAAGAGAGAGCTGTGGGAGGAGAAAGACCGTTGGAGTGAACACAGAGTTCTG 1380
Db |||||||
QY 1321 AGAAGAGAGAGCTGTGGGAGGAGAAAGACCGTTGGAGTGAACACAGAGTTCTG 1380
Db |||||||
QY 1381 GCAGAGACAAAGAACTATACCTCAATCAAGGAAAAAGAGCCGAGGAGACCATGCTCTGACA 1440
Db |||||||
QY 1381 GCAGAGACAAAGAACTATACCTCAATCAAGGAAAAAGAGCCGAGGAGACCATGCTCTGACA 1440
Db |||||||
QY 1441 AGAATGAGGAGGAGAAAGAAAGAACTTTTAAAGCTGTGTGGATTGATGCACTCATTCAG 1500
Db |||||||
QY 1441 AGAATGAGGAGGAGAAAGAAAGAACTTTTAAAGCTGTGTGGATTGATGCACTCATTCAG 1500
Db |||||||
QY 1501 AAAAATACTACTCTCAGTGAACCCATGGATCAGGTGGGAGATTCTACAGTGTGGAAACGA 1560
Db |||||||
QY 1501 AAAAATACTACTCTCAGTGAACCCATGGATCAGGTGGGAGATTCTACAGTGTGGAAACGA 1560
Db |||||||

QY 1561 GTAGGCTTCGTGACTTATATGACAAATTTTGAGGAGGAGTTGGGAGCAGCGCAAGAAAGG 1620
DB |||||
QY 1561 GTAGGCTTCGTGACTTATATGACAAATTTTGAGGAGGAGTTGGGAGCAGCGCAAGAAAGG 1620
DB |||||
QY 1621 CCAAAGCTGCTCGGCCCTCCGTGGGAACTCCAAAGAGCAGAGCTCGATGAAGATTATAGAGA 1680
DB |||||
QY 1621 CCAAAGCTGCTCGGCCCTCCGTGGGAACTCCAAAGAGCAGAGCTCGATGAAGATTATAGAGA 1680
DB |||||
QY 1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGAGCAGACCTGTTCTAGCAGCTCAGACT 1740
DB |||||
QY 1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGAGCAGACCTGTTCTAGCAGCTCAGACT 1740
DB |||||
QY 1741 CTGAAGTTTTTGACGTTATTGTCAGAAATCAAAACGCAAAAGGCCACCCCTGACCGACTTC 1800
DB |||||
QY 1741 CTGAAGTTTTTGACGTTATTGTCAGAAATCAAAACGCAAAAGGCCACCCCTGACCGACTTC 1800
DB |||||
QY 1801 ATGATGAACCTTTGGTCAACAGATCCAGGCCAGATGAATGATGACACACTCTGCAAAATGCA 1860
DB |||||
QY 1801 ATGATGAACCTTTGGTCAACAGATCCAGGCCAGATGAATGATGACACACTCTGCAAAATGCA 1860
DB |||||
QY 1861 GGCAAAGGCAAGACGACAGGAATTTAGGCAAGCAATTTATCTGAGAGAGGCCATCA 1920
DB |||||
QY 1861 GGCAAAGGCAAGACGACAGGAATTTAGGCAAGCAATTTATCTGAGAGAGGCCATCA 1920
DB |||||
QY 1921 AGCCCTGCTGCTATGACCAACATGCTGGCAGACTTTTCCACTACCGGATCACAGTCT 1980
DB |||||
QY 1921 AGCCCTGCTGCTATGACCAACATGCTGGCAGACTTTTCCACTACCGGATCACAGTCT 1980
DB |||||
QY 1981 CCCGCTGACGAATTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCAGAGT 2040
DB |||||
QY 1981 CCCGCTGACGAATTTTAACTGACAGGCCAACTGTTATAGAAATACGATGATCAGAGT 2040
DB |||||
QY 2041 ATATCTTTGAAGGATTTCTATGTTTGACATGCCCTCCCTGACCAATATTTCCACTGTGTA 2100
DB |||||
QY 2101 AAGTAATTAGATTTCAACATAGACTACAGATTCATTTCAATTTGAAGAGATGATGCCGAGA 2160
DB |||||
QY 2101 AAGTAATTAGATTTCAACATAGACTACAGATTCATTTCAATTTGAAGAGATGATGCCGAGA 2160
DB |||||
QY 2161 ATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTCTTATTCAGAGATATTTTGAAT 2220
DB |||||
QY 2161 ATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTCTTATTCAGAGATATTTTGAAT 2220
DB |||||
QY 2221 TATATGACTGGAATCTTAAAGGTCCTTTGTTGAAGACAGCCCTCCCTGCTGCCAAGAT 2280
DB |||||
QY 2221 TATATGACTGGAATCTTAAAGGTCCTTTGTTGAAGACAGCCCTCCCTGCTGCCAAGAT 2280
DB |||||
QY 2281 TTCAATTTCAATGCCACGTTTGAAGATTTCTTCCAGATGAGGAAAGGAGTCTGTCCA 2340
DB |||||
QY 2281 TTCAATTTCAATGCCACGTTTGAAGATTTCTTCCAGATGAGGAAAGGAGTCTGTCCA 2340
DB |||||
QY 2341 TGCACAGATTTCTCTGTACTTGTAAAGTGTGACAGAAAGCCCTGCTGAGGAGGAGA 2400
DB |||||
QY 2341 TGCACAGATTTCTCTGTACTTGTAAAGTGTGACAGAAAGCCCTGCTGAGGAGGAGA 2400
DB |||||
QY 2401 TTGCCAATATGCTTCAGTGGGAGAGCTGGAGTGGCAGAAAATATGACAGAAATGCAAAAG 2460
DB |||||
QY 2401 TTGCCAATATGCTTCAGTGGGAGAGCTGGAGTGGCAGAAAATATGACAGAAATGCAAAAG 2460
DB |||||
QY 2461 GCATGATTTGTTACCAACCCCTGGGAGCAAAACCAAGCTCTGTCCGTAACGATCAACTGGATC 2520
DB |||||
QY 2461 GCATGATTTGTTACCAACCCCTGGGAGCAAAACCAAGCTCTGTCCGTAACGATCAACTGGATC 2520
DB |||||
QY 2521 GTGAAACAGTTCAACCCGATGTGATTTCTTCCGATTTATCTGTCACCTTTGGGATACGCC 2580
DB |||||
QY 2521 GTGAAACAGTTCAACCCGATGTGATTTCTTCCGATTTATCTGTCACCTTTGGGATACGCC 2580
DB |||||
QY 2581 CTGCACAGTTGAGTTATGACAGGAGACCCACAGTACCAAAACTGTGGAAGAGTTATGTGA 2640
DB |||||
QY 2581 CTGCACAGTTGAGTTATGACAGGAGACCCACAGTACCAAAACTGTGGAAGAGTTATGTGA 2640
DB |||||
QY 2641 AACTTCCGCCACCTCTTAGCAAAATAGTCCCAAGTCAAACTGACAAACAGAGAGCTGG 2700

DB |||||
QY 2641 AACTTCCGCCACCTCTTAGCAAAATAGTCCCAAGTCAAACTGACAAACAGAGAGCTGG 2700
DB |||||
QY 2701 CACAGAGGAGGAGAGCCCTCCAAAAATACGGCAGAGAAATACAATGACAGAGAGTAA 2760
DB |||||
QY 2701 CACAGAGGAGGAGAGCCCTCCAAAAATACGGCAGAGAAATACAATGACAGAGAGTAA 2760
DB |||||
QY 2761 CGGTGAGCTTAAGTAGCAGAGGATTTCTGGAACACTGGCATCCCTTCTGATCTGTGTCAGC 2820
DB |||||
QY 2761 CGGTGAGCTTAAGTAGCAGAGGATTTCTGGAACACTGGCATCCCTTCTGATCTGTGTCAGC 2820
DB |||||
QY 2821 ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCTAATGCATT 2880
DB |||||
QY 2821 ATGCAATGATGCTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCTAATGCATT 2880
DB |||||
QY 2881 TGAACAAGTTGATAGGATATATCTTTCCAAAGATGCTGTTCTGTTGAGCTGGGCCATGACTC 2940
DB |||||
QY 2881 TGAACAAGTTGATAGGATATATCTTTCCAAAGATGCTGTTCTGTTGAGCTGGGCCATGACTC 2940
DB |||||
QY 2941 ATCCAGTCTCATCTTTAAATTTTGGAAATGAAATCTCTGATCATGCGCAAGATTCATATCTA 3000
DB |||||
QY 2941 ATCCAGTCTCATCTTTAAATTTTGGAAATGAAATCTCTGATCATGCGCAAGATTCATATCTA 3000
DB |||||
QY 3001 ACTGTGGAATTCGGCAGCCAAATACGGAGACAGAGAAAGTTTCAATCATGACATGCAATCGGA 3060
DB |||||
QY 3001 ACTGTGGAATTCGGCAGCCAAATACGGAGACAGAGAAAGTTTCAATCATGACATGCAATCGGA 3060
DB |||||
QY 3061 AGAAAGGATTAACACCTTTGATAAATATCATGTGACGCCCTTGGCCCAAGATGACCCAACTC 3120
DB |||||
QY 3061 AGAAAGGATTAACACCTTTGATAAATATCATGTGACGCCCTTGGCCCAAGATGACCCAACTC 3120
DB |||||
QY 3121 CCTCGAGGATTAACCAATGAAATGAAATTTCTGGGTGATGCTGTTGTTGAATTTTC 3180
DB |||||
QY 3121 CCTCGAGGATTAACCAATGAAATGAAATTTCTGGGTGATGCTGTTGTTGAATTTTC 3180
DB |||||
QY 3181 TGAACAGCTTCATTTGTACTATTTCTTCTAGTCTGGAAGAGGAGGATTAAGCAACCT 3240
DB |||||
QY 3181 TGAACAGCTTCATTTGTACTATTTCTTCTAGTCTGGAAGAGGAGGATTAAGCAACCT 3240
DB |||||
QY 3241 ATCGGACTGCCATTTGTTGAGAATCAGACACTTGGCCATGCTAGCAAGAAACTTGAACCTGG 3300
DB |||||
QY 3241 ATCGGACTGCCATTTGTTGAGAATCAGACACTTGGCCATGCTAGCAAGAAACTTGAACCTGG 3300
DB |||||
QY 3301 ATCCATTTATGCTGTATGCTCAGGCCCTGACCTTTGTAGAGAACTCGGACCTTCGACATG 3360
DB |||||
QY 3301 ATCCATTTATGCTGTATGCTCAGGCCCTGACCTTTGTAGAGAACTCGGACCTTCGACATG 3360
DB |||||
QY 3361 CAATGSCCAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGAGGAGCTGGAGG 3420
DB |||||
QY 3361 CAATGSCCAATTTGTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGAGGAGCTGGAGG 3420
DB |||||
QY 3421 AAGCCCAAGCAGTTTATTTGGACGCTTGTCTTTTAAATGATCCGGAACCTTCGCGAAGTCTGGC 3480
DB |||||
QY 3421 AAGCCCAAGCAGTTTATTTGGACGCTTGTCTTTTAAATGATCCGGAACCTTCGCGAAGTCTGGC 3480
DB |||||
QY 3481 TCAATTTATCTCTCCACCCACTCCAACACTAAGAGCAAAATCTGATCGACAACTTATTTG 3540
DB |||||
QY 3481 TCAATTTATCTCTCCACCCACTCCAACACTAAGAGCAAAATCTGATCGACAACTTATTTG 3540
DB |||||
QY 3541 AAACCTTCTCCAGTTCTTACAAAACCTTACTGAGTTTGAAGCAATTTGGAGTAATTTTAA 3600
DB |||||
QY 3541 AAACCTTCTCCAGTTCTTACAAAACCTTACTGAGTTTGAAGCAATTTGGAGTAATTTTAA 3600
DB |||||
QY 3601 CTCATGTTTCGACTTCTGGCAAGGCAATTCATTTGAGAACTGTGGGATTTTAAACCATCTGA 3660
DB |||||
QY 3601 CTCATGTTTCGACTTCTGGCAAGGCAATTCATTTGAGAACTGTGGGATTTTAAACCATCTGA 3660
DB |||||
QY 3661 CCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATTAATGCAACTGGTAGCCA 3720
DB |||||
QY 3661 CCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGACTCCATTAATGCAACTGGTAGCCA 3720
DB |||||
QY 3721 CAGAGTACTTATTCATTTTCCAGATCATCATGAGGACACTTAACTTTGTTGTCGGA 3780

Db 3721 CAGAGTACTTATTTCAATTCATTTCCAGATCATCATGAAGGACACTTAACTTTGTCGAA 3780
Qy 3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATCGCAGACT 3840
Db 3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATCGCAGACT 3840
Qy 3841 ACGCCATAACCAACGACAGACCAAGAGGCGCTGTGGCGCTTCGCCAACAGACCTTGGCGG 3900
Db 3841 ACGCCATAACCAACGACAGACCAAGAGGCGCTGTGGCGCTTCGCCAACAGACCTTGGCGG 3900
Qy 3901 ACCTTTTGAATCATTTATTTGAGCGCTGTGACACTGATGAAGATTTGGAATATGTTTCATA 3960
Db 3901 ACCTTTTGAATCATTTATTTGACGCTGTGACACTGATGAAGATTTGGAATATGTTTCATA 3960
Qy 3961 CTTTCATGATGCTCTCTTTTCCAGATTTGAAGAAATTCATTTTGAATCAGGATTTGA 4020
Db 3961 CTTTCATGATGCTCTCTTTTCCAGATTTGAAGAAATTCATTTTGAATCAGGATTTGA 4020
Qy 4021 ATGACCCCAAATCCAGCTTCAGCAGTGTGCTTGACACTTAGGACAGAGGAAAGAGC 4080
Db 4021 ATGACCCCAAATCCAGCTTCAGCAGTGTGCTTGACACTTAGGACAGAGGAAAGAGC 4080
Qy 4081 CAGACATTCCTCTGTACAGACTCTCGACAGTGGGCGCATCCCATGCGCGAACCTACA 4140
Db 4081 CAGACATTCCTCTGTACAGACTCTCGACAGTGGGCGCATCCCATGCGCGAACCTACA 4140
Qy 4141 CTGTGCTGTATTTCAAGGGAAGAAATAGGCTGTGGGAAGGACCAAGTATTCAGC 4200
Db 4141 CTGTGCTGTATTTCAAGGGAAGAAATAGGCTGTGGGAAGGACCAAGTATTCAGC 4200
Qy 4201 AAGCGAAATGGGAGCAGCAATGGATGCGCTTGAAAAATATATTTTCCAGATGGCCC 4260
Db 4201 AAGCGAAATGGGAGCAGCAATGGATGCGCTTGAAAAATATATTTTCCAGATGGCCC 4260
Qy 4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTAAAGAAATGAGTGGGAAA 4320
Db 4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTTAAAGAAATGAGTGGGAAA 4320
Qy 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAATTAAGAGGGCAT 4380
Db 4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGAGACATCAAGAATTAAGAGGGCAT 4380
Qy 4381 GCAAGTGTGGAGTATTTACTTGCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
Db 4381 GCAAGTGTGGAGTATTTACTTGCTCAGTAACTGTGACTGTGTCTATTGAGACCTAGCCT 4440
Qy 4441 AGTTTTCTGCAGACAATGAAACGAACTGTCTCAITGAAATAAATACAGAGTCAAATCG 4500
Db 4441 AGTTTTCTGCAGACAATGAAACGAACTGTCTCAITGAAATAAATACAGAGTCAAATCG 4500
Qy 4501 CTATTGTTGTTTAAATGATCTGTTTTAGTGGATGGTCTTTATTTACAAAGTATTAGATT 4560
Db 4501 CTATTGTTGTTTAAATGATCTGTTTTAGTGGATGGTCTTTATTTACAAAGTATTAGATT 4560
Qy 4561 TTTCTCTATTTAAACGGAACCTTGACTTTGGTGAATGTCATTCTCTCTTTATTTTG 4620
Db 4561 TTTCTCTATTTAAACGGAACCTTGACTTTGGTGAATGTCATTCTCTCTTTATTTTG 4620
Qy 4621 CTCTTTTAAATAATAAATCAAGAGCATATTTCTATGTGGAATAGATCTGTTTTTCCAT 4680
Db 4621 CTCTTTTAAATAATAAATCAAGAGCATATTTCTATGTGGAATAGATCTGTTTTTCCAT 4680
Qy 4681 CTGTGTCAGATGTGACCTTAGACTTTCAATTTGACAGTAAATAATTTGACTTTACTAG 4740
Db 4681 CTGTGTCAGATGTGACCTTAGACTTTCAATTTGACAGTAAATAATTTGACTTTACTAG 4740
Qy 4741 TAAAAAATTT 4764
Db 4741 TAAAAAATTT 4764

LOCUS CS042637 4764 bp DNA linear PAT 22-MAR-2005
DEFINITION Sequence 3191 from Patent WO2005019258.
ACCESSION CS042637
VERSION CS042637.1 GI:61849798
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Abbas, A., Clark, H., Ouyang, W., Williams, P.M., Wood, W.I. and Wu, T.D.
TITLE Compositions and methods for the treatment of immune related diseases
JOURNAL Patent: WO 2005019258-A 3191 03-MAR-2005;
Genentech, Inc. (US)
FEATURES
source
1..4764
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.9%; Score 4760.8; DB 6; Length 4764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CTGCTCTGTGACTCTGGGTAGTAGCCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG 60
Db 1 CTGCTCTGTGACTCTGGGTAGTAGCCTGGCTTTGCTCTGACGGCGATCTCGGGCCCGAG 60
Qy 61 AGCCTTTTATAGTGTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
Db 61 AGCCTTTTATAGTGTGCTTTTCCCGGGATGTGAAGGATACAGAAATGACTGTGAATCAA 120
Qy 121 CCCATATCATCAAGGAGCTGATAATCTAGTGAAGAGTTAGACGTGTGCATCTTCACTA 180
Db 121 CCCATATCATCAAGGAGCTGATAATCTAGTGAAGAGTTAGACGTGTGCATCTTCACTA 180
Qy 181 TGATATGAGGAGCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCGGA 240
Db 181 TGATATGAGGAGCTCTGAGCTTATATCTCTGTGGAAGATGTGACATATCCAGGCGGA 240
Qy 241 ACATCATGATGACAGGAAACACATGTCAAGAAATGTGTTCCACCGGAGCAGAGGCGTC 300
Db 241 ACATCATGATGACAGGAAACACATGTCAAGAAATGTGTTCCACCGGAGCAGAGGCGTC 300
Qy 301 CCCGAGGACGAGGAGGACATGGAGCCAGACCTCTCAGCACCATCTCTTTAGGCCCCAAAATC 360
Db 301 CCCGAGGACGAGGAGGACATGGAGCCAGACCTCTCAGCACCATCTCTTTAGGCCCCAAAATC 360
Qy 361 TGAGGCTGTTCACTCTCAGGAGCTCTGTCGATATCAATATGACCTCCAAGTGCCC 420
Db 361 TGAGGCTGTTCACTCTCAGGAGCTCTGTCGATATCAATATGACCTCCAAGTGCCC 420
Qy 421 CTTCCACCACTTTCTCAAACTCTCAGAGCCCCCAATTTCTCCCTCCAGCAGCAGACTTTG 480
Db 421 CTTCCACCACTTTCTCAAACTCTCAGAGCCCCCAATTTCTCCCTCCAGCAGCAGACTTTG 480
Qy 481 TACCTTTCCCCCACCACCATGCTCGGTCAAGCAGAGGCGCTCTTCCCTCTGCCCAATCA 540
Db 481 TACCTTTCCCCCACCACCATGCTCGGTCAAGCAGAGGCGCTCTTCCCTCTGCCCAATCA 540
Qy 541 GGCAGGCTTTTCCCAACCAACAGATGAGGACACCCCTTCCAGTTCCTCTGTTTTCCTC 600
Db 541 GGCAGGCTTTTCCCAACCAACAGATGAGGACACCCCTTCCAGTTCCTCTGTTTTCCTC 600
Qy 601 CCATGCCACCAATGCTCTGCTTAATAACCCCTCAGTCCCTGGGGCACCTCTCTGGAC 660
Db 601 CCATGCCACCAATGCTCTGCTTAATAACCCCTCAGTCCCTGGGGCACCTCTCTGGAC 660
Qy 661 AAGGCACTTTTCCCTTTCATGATGCCCCCTCCCTCCATGCTCATCCCCCGCCCCCTCCAG 720
Db 661 AAGGCACTTTTCCCTTTCATGATGCCCCCTCCCTCCATGCTCATCCCCCGCCCCCTCCAG 720

Db 661 AAGGCACTTTCCCTTTCATGATGCCCTCCCTCCATGCTCATCCCCCGCCCCCTCCAG 720
Qy 721 TCATGCCGAGCAGGTTAATATATACGTACCTCCGGGCTATTCTCACCAACATTTCCAC 780
Db 721 TCATGCCGAGCAGGTTAATATATACGTACCTCCGGGCTATTCTCACCAACATTTCCAC 780
Qy 781 CTCCTCAGTTTAAATAGTTTCCAGAACCAACCTAGTTCTTTCTCTGCCCCAGTGTATAACA 840
Db 781 CTCCTCAGTTTAAATAGTTTCCAGAACCAACCTAGTTCTTTCTCTGCCCCAGTGTATAACA 840
Qy 841 GCAGTAGTCTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAAAGGCTCCCAAGTGAGA 900
Db 841 GCAGTAGTCTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAAAGGCTCCCAAGTGAGA 900
Qy 901 GAAAGTCCCAAGAAAGGCTGAAACATATGATGACCAACAGGACCCAGACCAAGTCATG 960
Db 901 GAAAGTCCCAAGAAAGGCTGAAACATATGATGACCAACAGGACCCAGACCAAGTCATG 960
Qy 961 GCGAGGTGAGAGGATCGGTCCCTGGATCGGGGAGCGAGCCGACATCCCGACAGGA 1020
Db 961 GCGAGGTGAGAGGATCGGTCCCTGGATCGGGGAGCGAGCCGACATCCCGACAGGA 1020
Qy 1021 GAAAGCAAGACAGCCGGTACAGATCTGATTTATGACCGAGGAGAACCAACATCTCGCCACC 1080
Db 1021 GAAAGCAAGACAGCCGGTACAGATCTGATTTATGACCGAGGAGAACCAACATCTCGCCACC 1080
Qy 1081 GCAGCTACGAAACGAGCAGAGCGAGAACCGGAGAGACACAGGACATCGAGCAACCGGAA 1140
Db 1081 GCAGCTACGAAACGAGCAGAGCGAGAACCGGAGAGACACAGGACATCGAGCAACCGGAA 1140
Qy 1141 GATCACCATCTCTGGAAGGCTCTACAAAAGAGTATAGAGATCTGGAAGGAGTTACG 1200
Db 1141 GATCACCATCTCTGGAAGGCTCTACAAAAGAGTATAGAGATCTGGAAGGAGTTACG 1200
Qy 1201 GTTTATCGGTGTCTGACCTCTGAGTGCACACAGAAATACCTCGGGAGATTATTA 1260
Db 1201 GTTTATCGGTGTCTGACCTCTGAGTGCACACAGAAATACCTCGGGAGATTATTA 1260
Qy 1261 AAAATAACAGATCTTTGGGGCCCAACCTCTGGAGATTGTGAATCATCGCTCCCAAGTAGGG 1320
Db 1261 AAAATAACAGATCTTTGGGGCCCAACCTCTGGAGATTGTGAATCATCGCTCCCAAGTAGGG 1320
Qy 1321 AGAAGAGAGAGCTCTTTGGGAGGAGAAAGAACCGTTGGAGTGACAAACAGAGTTCTG 1380
Db 1321 AGAAGAGAGAGCTCTTTGGGAGGAGAAAGAACCGTTGGAGTGACAAACAGAGTTCTG 1380
Qy 1381 GCAAGACAGAAATATATACCTCAATCAAGGAAAGAGCCCGAGGAGACCATGCTGACA 1440
Db 1381 GCAAGACAGAAATATATACCTCAATCAAGGAAAGAGCCCGAGGAGACCATGCTGACA 1440
Qy 1441 AGAATGAGGAGGAGAAAGAAAGAACTTTTAAAGCTGTGTGGATTGATGCATCTCATTTCA 1500
Db 1441 AGAATGAGGAGGAGAAAGAAAGAACTTTTAAAGCTGTGTGGATTGATGCATCTCATTTCA 1500
Qy 1501 AAAAATCTACTCTCAGTGACCCCATGGATCAGTGGGAGATTCTACAGTGGTTGGAAAG 1560
Db 1501 AAAAATCTACTCTCAGTGACCCCATGGATCAGTGGGAGATTCTACAGTGGTTGGAAAG 1560
Qy 1561 GTAGGCTCTGATCTATATGACAAATTTGAGGAGGAGTTGGGAGCAGGCAAGAAAGG 1620
Db 1561 GTAGGCTCTGATCTATATGACAAATTTGAGGAGGAGTTGGGAGCAGGCAAGAAAGG 1620
Qy 1621 CCAAGAGCTCTCGGCTCCGTGGGAACCTCCAAAGACGAAGTCCGATGAAGATTATAGAG 1680
Db 1621 CCAAGAGCTCTCGGCTCCGTGGGAACCTCCAAAGACGAAGTCCGATGAAGATTATAGAG 1680
Qy 1681 GTTCCAGTGAATCCGAGTGTGAGTGTGATGAGGACAGCACTGTTCTAGCAGTCAAGCT 1740
Db 1681 GTTCCAGTGAATCCGAGTGTGAGTGTGATGAGGACAGCACTGTTCTAGCAGTCAAGCT 1740
Qy 1741 CTGAGTTTTTGCAGTTATTCGAGAAATCAACCGGAAAGGCCCCACCTGACGACTTC 1800
Db 1741 CTGAGTTTTTGCAGTTATTCGAGAAATCAACCGGAAAGGCCCCACCTGACGACTTC 1800

Qy 1801 ATGATGAACCTTTGGTACAACGATCCAGGCCAGATGAATGAATGATGACCACTCTGCAATGCA 1860
Db 1801 ATGATGAACCTTTGGTACAACGATCCAGGCCAGATGAATGAATGATGACCACTCTGCAATGCA 1860
Qy 1861 GGGCAAGGCAAGAGCGCACAGGAAATTTAGGCACAGCATTTTATCTCTGAGAAAGAGGCCATCA 1920
Db 1861 GGGCAAGGCAAGAGCGCACAGGAAATTTAGGCACAGCATTTTATCTCTGAGAAAGAGGCCATCA 1920
Qy 1921 AGCCCTGTCTCTATGACCAACAAATGTCTGGCAGACTTTTCCACTACCGGATCACAGTCT 1980
Db 1921 AGCCCTGTCTCTATGACCAACAAATGTCTGGCAGACTTTTCCACTACCGGATCACAGTCT 1980
Qy 1981 CCGCGCTACGAACTTTTAACTGACAGGCGCAACTGTGTATAGAAATACGATGATCAGAGT 2040
Db 1981 CCGCGCTACGAACTTTTAACTGACAGGCGCAACTGTGTATAGAAATACGATGATCAGAGT 2040
Qy 2041 ATATCTTTGAAAGGATTTTCTATGTTTGCACATGCCCCCTCGACCAATATTTCCACTGTGTA 2100
Db 2041 ATATCTTTGAAAGGATTTTCTATGTTTGCACATGCCCCCTCGACCAATATTTCCACTGTGTA 2100
Qy 2101 AAGTAATTAGATTCAACATAGACTACAGATTCATTTTCAATGGAAGAGATGATGCCGAGA 2160
Db 2101 AAGTAATTAGATTCAACATAGACTACAGATTCATTTTCAATGGAAGAGATGATGCCGAGA 2160
Qy 2161 ATTTTGTGTGAAAGGCTTGAACCTCTTTCACTGTTCCTATTTCCAGATATTTTGGAAAT 2220
Db 2161 ATTTTGTGTGAAAGGCTTGAACCTCTTTCACTGTTCCTATTTCCAGATATTTTGGAAAT 2220
Qy 2221 TATATGACTGGAATCTTAAAGGTCTTTGTTTGAAGACAGCCCTCTCTGCTGCCCAAGAT 2280
Db 2221 TATATGACTGGAATCTTAAAGGTCTTTGTTTGAAGACAGCCCTCTCTGCTGCCCAAGAT 2280
Qy 2281 TTCAATTTCAATGACAGCTTTTGAAGATTTCTTCCAGATGGAGGAAAGGAGTCTGTCCA 2340
Db 2281 TTCAATTTCAATGACAGCTTTTGAAGATTTCTTCCAGATGGAGGAAAGGAGTCTGTCCA 2340
Qy 2341 TGCACAGATTTCTCTGTACTTTTAAAGGTGAGTGCAGCAAGCCCTGTGTCTGAGGAGGAGA 2400
Db 2341 TGCACAGATTTCTCTGTACTTTTAAAGGTGAGTGCAGCAAGCCCTGTGTCTGAGGAGGAGA 2400
Qy 2401 TTGGCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGACAGAAAGATGCAAG 2460
Db 2401 TTGGCAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGACAGAAAGATGCAAG 2460
Qy 2461 GCATGATTTGTTACCAACCCCTGGGACGAAACCAAGCTCTCTCGTATCGATCAACTGGATC 2520
Db 2461 GCATGATTTGTTACCAACCCCTGGGACGAAACCAAGCTCTCTCGTATCGATCAACTGGATC 2520
Qy 2521 GTGAACAGTTTCAACCCCGATGTGATTTATTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580
Db 2521 GTGAACAGTTTCAACCCCGATGTGATTTATTTTCCGATTTATCGTCCACTTTGGGATACGCC 2580
Qy 2581 CTGCAAGTTGAGTTATGAGGAGACCCACAGTACCAAACTGTGGAAGAGTTATGTGA 2640
Db 2581 CTGCAAGTTGAGTTATGAGGAGACCCACAGTACCAAACTGTGGAAGAGTTATGTGA 2640
Qy 2641 AACTTCCGACCTCTCTAGCAAAATAGTCCCAAAAGTCAAAACAACTGCAAAACGAGAGCTGG 2700
Db 2641 AACTTCCGACCTCTCTAGCAAAATAGTCCCAAAAGTCAAAACAACTGCAAAACGAGAGCTGG 2700
Qy 2701 CACAGAGGAGGAGAACCCCTCCAAAATAAGTCCCAAAAGTCAAAACAACTGCAAAACGAGAGTAA 2760
Db 2701 CACAGAGGAGGAGAACCCCTCCAAAATAAGTCCCAAAAGTCAAAACAACTGCAAAACGAGAGTAA 2760
Qy 2761 CCGTGGAGTAAATAGCAGGAGTCTGGAAACTGGGATCCGTTCTGATGCTGTGTCAGC 2820
Db 2761 CCGTGGAGTAAATAGCAGGAGTCTGGAAACTGGGATCCGTTCTGATGCTGTGTCAGC 2820
Qy 2821 ATGCAATGATGCTACTCTGTTCTGACCCATCATATCGCTTACCAACCAATGCTTAATGCAAT 2880
Db 2821 ATGCAATGATGCTACTCTGTTCTGACCCATCATATCGCTTACCAACCAATGCTTAATGCAAT 2880

QY 2881 TGGCAAGATTGATAGGATATACCTTTCCAGATCGTTGTCTGTGTGAGCTGGCCATGACTC 2940
DB 2881 TGGCAAGATTGATAGGATATACCTTTCCAGATCGTTGTCTGTGTGAGCTGGCCATGACTC 2940
QY 2941 ATCCAAAGTCATCATTTTAAATTTTGGAAATGAAATCTGTATCATGCCAGGAATTCATTTATCTA 3000
DB 2941 ATCCAAAGTCATCATTTTAAATTTTGGAAATGAAATCTGTATCATGCCAGGAATTCATTTATCTA 3000
QY 3001 ACTGTGGAATTCGGCAGCCCAATACGGGAGACAGAAAAGTTTCATCACATGCACATGCGGA 3060
DB 3001 ACTGTGGAATTCGGCAGCCCAATACGGGAGACAGAAAAGTTTCATCACATGCACATGCGGA 3060
QY 3061 AGAAGGGGAAATAACACTTTGATAAATATCATGTCAAGCTTGGCCCAAGATGACCCAACTC 3120
DB 3061 AGAAGGGGAAATAACACTTTGATAAATATCATGTCAAGCTTGGCCCAAGATGACCCAACTC 3120
QY 3121 CCTCGAGGAAATAACCAATGAAACGGTTGGAATTCCTGGGTGATGCTGTGTGTGAATTTTC 3180
DB 3121 CCTCGAGGAAATAACCAATGAAACGGTTGGAATTCCTGGGTGATGCTGTGTGTGAATTTTC 3180
QY 3181 TGACCAAGCTCCATTTGTACTATTTGTTTCTAGTCTGGAAGAGGAGATTTAGCAACCT 3240
DB 3181 TGACCAAGCTCCATTTGTACTATTTGTTTCTAGTCTGGAAGAGGAGATTTAGCAACCT 3240
QY 3241 ATCGGACTGCAATGTTTCAGAAATCAGACCTTGGCCATGCTAGCAAGAAATTTGAACCTGG 3300
DB 3241 ATCGGACTGCAATGTTTCAGAAATCAGACCTTGGCCATGCTAGCAAGAAATTTGAACCTGG 3300
QY 3301 ATCCATTTATGCTGTATGCTCAAGGCTTGAACCTTTGTGAGAAATCGGACCTTGGACATG 3360
DB 3301 ATCCATTTATGCTGTATGCTCAAGGCTTGAACCTTTGTGAGAAATCGGACCTTGGACATG 3360
QY 3361 CAATGCCCAATGTTTGAAGCTTAATAGGAGCTGTTTACTTGGAGGAAAGCTTGGAGG 3420
DB 3361 CAATGCCCAATGTTTGAAGCTTAATAGGAGCTGTTTACTTGGAGGAAAGCTTGGAGG 3420
QY 3421 AAGCCAAAGCAGTTATTTGGAAGCTTCTCTTTTAAATGATCGGACCTTGGCAAGTCTGGC 3480
DB 3421 AAGCCAAAGCAGTTATTTGGAAGCTTCTCTTTTAAATGATCGGACCTTGGCAAGTCTGGC 3480
QY 3481 TCAATTTATCCTCTCCACCACTCCAACTCAAGAGCCAAATCTGATGACCAACTTTATTTG 3540
DB 3481 TCAATTTATCCTCTCCACCACTCCAACTCAAGAGCCAAATCTGATGACCAACTTTATTTG 3540
QY 3541 AACTCTCTCAGTTCTACAAAACCTTACTGAGTTTGAAGAACAAATGGAATTTTGA 3600
DB 3541 AACTCTCTCAGTTCTACAAAACCTTACTGAGTTTGAAGAACAAATGGAATTTTGA 3600
QY 3601 CTCAATGTTGACCTTCTGGCAAGGCAATTCATTTGAGAACTGTGGCAATTTAACCATCTGA 3660
DB 3601 CTCAATGTTGACCTTCTGGCAAGGCAATTCATTTGAGAACTGTGGCAATTTAACCATCTGA 3660
QY 3661 CCTTAGGCCCAATTCAGAGAAATGGAATTCCTAGGTGATCTCAATATGCAACTGGTAGCCA 3720
DB 3661 CCTTAGGCCCAATTCAGAGAAATGGAATTCCTAGGTGATCTCAATATGCAACTGGTAGCCA 3720
QY 3721 CAGAGTACTTTATTCATTTCCAGATCATCATGAGGACCTTTAACTTTGTTCGGA 3780
DB 3721 CAGAGTACTTTATTCATTTCCAGATCATCATGAGGACCTTTAACTTTGTTCGGA 3780
QY 3781 GCTCTTTGGTGAATATAGAACTCAGGCCAAGTGTAGCGGAGGAGCTGGGCATGCAGAGT 3840
DB 3781 GCTCTTTGGTGAATATAGAACTCAGGCCAAGTGTAGCGGAGGAGCTGGGCATGCAGAGT 3840
QY 3841 ACGCCATAACCAACGACCAAGAGGCTGTGGCGCTTCGCAACCAAGACCTTGGCGG 3900
DB 3841 ACGCCATAACCAACGACCAAGAGGCTGTGGCGCTTCGCAACCAAGACCTTGGCGG 3900
QY 3901 ACCTTTTGGAAATCATTTATTTGCGCGCTGTACATCTGATAGGAATTTGGAAATATGTTTCA 3960
DB 3901 ACCTTTTGGAAATCATTTATTTGCGCGCTGTACATCTGATAGGAATTTGGAAATATGTTTCA 3960
QY 3961 CTTTCATGAATGCTGCTCTTTCCACGATTTGAAAGAAATTCATTTTGAATCAGGATTGGA 4020

DB 3961 CTTTCATGAATGCTGCTCTTTCCAGATTTGAAAGAAATTCATTTTGAATCAGGATTGGA 4020
QY 4021 ATGACCCCAATCCAGCTTCAGCAGTGTTCCTTGACACCTTAGGACAGAGGAAAGAGC 4080
DB 4021 ATGACCCCAATCCAGCTTCAGCAGTGTTCCTTGACACCTTAGGACAGAGGAAAGAGC 4080
QY 4081 CAGACATTCCTCTGTAACAAGCTCTGACAGACAGTGGGCCCCATCCCATGCCGAACTTACA 4140
DB 4081 CAGACATTCCTCTGTAACAAGCTCTGACAGACAGTGGGCCCCATCCCATGCCGAACTTACA 4140
QY 4141 CTGTGGCTGTTTATTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTCAGC 4200
DB 4141 CTGTGGCTGTTTATTTCAAGGAGAAAGAAATAGGCTGTGGGAAAGGACCAAGTATTCAGC 4200
QY 4201 AAGCGGAAATGGGACAGCAATGGATGCGCTTGAAAAAATAAATTTTCCACAGATGGGCC 4260
DB 4201 AAGCGGAAATGGGACAGCAATGGATGCGCTTGAAAAAATAAATTTTCCACAGATGGGCC 4260
QY 4261 ATCAGAAGCGGTTTCATCGAAACGAAAGTACAGACAAAGTATAAAGAAATGAGGTGGGAAA 4320
DB 4261 ATCAGAAGCGGTTTCATCGAAACGAAAGTACAGACAAAGTATAAAGAAATGAGGTGGGAAA 4320
QY 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGGAAGACATCAAGAAATAAAGAGGAGGCAT 4380
DB 4321 GAGAGCATCAAGAGAGAGAGCCAGATGAGACTGGAAGACATCAAGAAATAAAGAGGAGGCAT 4380
QY 4381 GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACCTGTGTCTTATTTGAGACCTTAGCCT 4440
DB 4381 GCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACCTGTGTCTTATTTGAGACCTTAGCCT 4440
QY 4441 AGTTTTTCTTCGACAGCAATGAAACGAAAGTGTGCTCAATGGAATAAATAFACAGATCAAAATCG 4500
DB 4441 AGTTTTTCTTCGACAGCAATGAAACGAAAGTGTGCTCAATGGAATAAATAFACAGATCAAAATCG 4500
QY 4501 CTATTTGTTGTTTAAATGATCTGTTTGTAGCTGGATGCTCTTTTATTTACAAAGTATTAGATT 4560
DB 4501 CTATTTGTTGTTTAAATGATCTGTTTGTAGCTGGATGCTCTTTTATTTACAAAGTATTAGATT 4560
QY 4561 TTTCTTCTATTTAAACGGAACCTGCACTTTGTGTGAATGTGCAATTAATCTTCTTTTATTTTG 4620
DB 4561 TTTCTTCTATTTAAACGGAACCTGCACTTTGTGTGAATGTGCAATTAATCTTCTTTTATTTTG 4620
QY 4621 CTCCTTTAAATAATAAATTTCAAGAGACATTTCTATGTGGAATAGATCCTGTTTTCAT 4680
DB 4621 CTCCTTTAAATAATAAATTTCAAGAGACATTTCTATGTGGAATAGATCCTGTTTTCAT 4680
QY 4681 CTGTGTCCAGATTTGTGACCCCTAGACCTTCAATTTGACAGTAAAAAATTTGACTTTACTAG 4740
DB 4681 CTGTGTCCAGATTTGTGACCCCTAGACCTTCAATTTGACAGTAAAAAATTTGACTTTACTAG 4740
QY 4741 TAAAAAATTT 4764
DB 4741 TAAAAAATTT 4764

RESULT 4
LOCUS AF189011 4764 bp mRNA linear PRI 06-FEB-2001
DEFINITION Homo sapiens ribonuclease III (RN3) mRNA, complete cds.
ACCESSION AF189011
VERSION AF189011.1 GI:8886721
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 4764)
Wu, H., Xu, H., Miraglia, L. J. and Crooke, S. T.
Human RNase III is a 160-kDa protein involved in preribosomal RNA
processing
J. Biol. Chem. 275 (47), 36957-36965 (2000)

PUBMED 10948199
 REFERENCE 2 (bases 1 to 4764)
 AUTHORS Xu, H., Wu, H., Miraglia, L. J. and Crooke, S. T.
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-1999) Structural Biology, ISIS Pharmaceuticals,
 2292 Faraday Avenue, Carlsbad, CA 92008, USA
 FEATURES
 Location/Qualifiers
 source 1..4764
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 gene 1..4764
 /gene="RN3"
 CDS 246..4370
 /gene="RN3"
 /function="involved in preribosomal RNA processing"
 /notes="RNase III"
 /codon_start=1
 /product="ribonuclease III"
 /protein_id="AAP80558.1"
 /db_xref="GI:8866722"
 /translation="MMQGNTHRMSFHPGRCPRGRGHGARPSPFRPQNLRLHP
 QQPPVOYVEPSASPTTSNSPAPNFPNLPDPFVPPMPSPSAQGLPPCPPIRPPF
 PNHMRHPVPVPCPPMPNPPVPGAPGQGTFFPMPPSPSPHPPPPPV
 PQOVNVOYPPGYSHNFPPPSPNSPNNPSSFLPSANNSSPHFPLPPPLKPSB
 RPSPELKHVDHRRHDSHGGRHSRLDRRGSRSPRRRQDSRYSRSDYDRGTPS
 RHRSYRSRERHRHNRHDSRSLRSYKKEYRSYSLSYSLVVPAPGTPPLP
 GHIINNDTSWAPPLEIVNHRSPSRKKRARWEEKDRWDSQNSQDKNYTSIKKBP
 EBTMPDKNEBEEELIKPWIRCTHSYENYSDDPMDQVGDSTVVGTSRLRDLDFEE
 ELGSRQEKAPKALPWPPEPKTLDLELSSSESDSDSTCSSSDSEVPDLAEI
 KRKQAPDLHDLWYNDPGQMDPLCKSAKARTGIRHSIYGEERAIKCRPWN
 NAGRLPHRYITVSPFNFLDPTVIEDDHEIFEGSFNFAHAPLNTNPLCKVIRN
 IDYTHIFBEMPNPCVKGLELFLFRDILEYLDNMLKGLPLEDSPPCPPRPHFM
 PRVPLPDPGKEVLSMHQILLYLRCSKALPVEEIANMLQWEELEWQYAEBCKM
 IVTNPKSPSVRIIDOLREQNPDPVTPPIIVHFGIRPAOLSYAGDPQYOKLWSYV
 KLHLLANSPKVKTDOKLAOREBALOKIROKNTWRREVTVELSSQGFMTGIRSDV
 COHMMPLVTHIRVHOCMLHDKLIGVTFODRCLLOLANTHPSHLNFGMNPDPAR
 NSLSNGCINQPKGRKQVHMMRKGINTLINMSRLQDDPTPSRINHNRLRFLG
 DAVSFLTSVHLIYFSPLEGGGLATYRTAIYVONHMLAKKLELDPMLYAHGPD
 CRESDLRHANCFEALIGAVYLEGLEAKQFGLRLLFNDPLREVWHLNPLHPLQL
 QSPNTDROLIETSPVLQKLEPEEAGVIFTHVRLIARATLRTVGFNHLNLTGHQNM
 EFLGDSIMOLVATEYLFIHFPDHHGLTLRSLVNNRTOAKVAEELGMOEYAITND
 KTKRPVALTKTLADLESFIALYTDKLEVYHTFMNVCFPPRLKEFLINDWDPK
 SLOQCCLTRTEGKEPDILYKLTQTVGSPSHARTTVAIVFKGRIGCGKGFSIQQA
 ENGAAMDALRYNFPQMAHQXRFIERKYRQELKEMRWEREHQREPDDEDIKK"
 ORIGIN
 Query Match 99.9%; Score 4760.8; DB 8; Length 4764;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4762; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CTGTCTTGTAAGTCTGGGTAGTACCTGGCTTGTCTGACGCGGATCTCGGGCCCGAG 60
 DB 1 CTGTCTTGTAAGTCTGGGTAGTACCTGGCTTGTCTGACGCGGATCTCGGGCCCGAG 60
 QY 61 AGCCTTTTATAGTCTCTTTCCCGGGATGTGAAGATACAGAAATGATGTGAATCAA 120
 DB 61 AGCCTTTTATAGTCTCTTTCCCGGGATGTGAAGATACAGAAATGATGTGAATCAA 120
 QY 121 CCCATATCATCAAGGAGCTGTAATCTAGTGAAGAGTTAGACGTGTCATCTTCACTA 180
 DB 121 CCCATATCATCAAGGAGCTGTAATCTAGTGAAGAGTTAGACGTGTCATCTTCACTA 180
 QY 181 TGATATGAGGAGTCTCTGAGCTTATATTCTCTGTGGAAGATGTGACATATCCAGCGGA 240
 DB 181 TGATATGAGGAGTCTCTGAGCTTATATTCTCTGTGGAAGATGTGACATATCCAGCGGA 240
 QY 241 ACATCATGATGAGGGAACACATGTGTCAGAAATGTCTGTCCACCGGACGAGGCGTC 300
 DB 241 ACATCATGATGAGGGAACACATGTGTCAGAAATGTCTGTCCACCGGACGAGGCGTC 300
 QY 301 CCGGAGGACGAGGAGACATGAGGCCAGACCCCTCAGCACCATCTTCTTAGGCCCCCAATC 360

301 CCGGAGGACGAGGAGACATGAGGCCAGACCCCTCAGCACCATCTCTTTAGGCCCCCAATC 360
 QY 361 TGAGGCTGCTTCAACCCCTCAGAGCCCTCTGTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
 DB 361 TGAGGCTGCTTCAACCCCTCAGAGCCCTCTGTGTGCAATATCAATATGAACCTTCAAGTGCCC 420
 QY 421 CTTTCCACCACTTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACGAGCTTTG 480
 DB 421 CTTTCCACCACTTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACGAGCTTTG 480
 QY 481 TACCTTTCCCCCACCACCATGCTCTCGTCAAGCGAAGGCCCTCTTCCCCCTGCCCCAATCA 540
 DB 481 TACCTTTCCCCCACCACCATGCTCTCGTCAAGCGAAGGCCCTCTTCCCCCTGCCCCAATCA 540
 QY 541 GGCCTGCTTTCCCAACCAACAGATGAGGACCCCTTCCAGTTCCTCTCTGTTTCTCTC 600
 DB 541 GGCCTGCTTTCCCAACCAACAGATGAGGACCCCTTCCAGTTCCTCTCTGTTTCTCTC 600
 QY 601 CCATGCCACCAATGCTTGTCTTAATACCCCTCAGTCCCTGGGGACCTTCTCTGGAC 660
 DB 601 CCATGCCACCAATGCTTGTCTTAATACCCCTCAGTCCCTGGGGACCTTCTCTGGAC 660
 QY 661 AAGGCACCTTTCCCTTTCATGATGCCCTTCCCTTCCATGCTCATCCCCCGCCCCCTCCAG 720
 DB 661 AAGGCACCTTTCCCTTTCATGATGCCCTTCCCTTCCATGCTCATCCCCCGCCCCCTCCAG 720
 QY 721 TCATGCCGAGCAGGTTAATATTCAGTACCTTCCGGGCTTATCTCACCAAACTTCCCAC 780
 DB 721 TCATGCCGAGCAGGTTAATATTCAGTACCTTCCGGGCTTATCTCACCAAACTTCCCAC 780
 QY 781 CTCCTCAGTTTAAATAGTTTCCAGAACACCCCTAGTCTTCTTCTGCCCCAGTGTCAATAACA 840
 DB 781 CTCCTCAGTTTAAATAGTTTCCAGAACACCCCTAGTCTTCTTCTGCCCCAGTGTCAATAACA 840
 QY 841 GCAGTAGTCTCATTTTCAGACATCTCCCTTCCATACCCATCTCCCAAGGCTCCCAAGTGA 900
 DB 841 GCAGTAGTCTCATTTTCAGACATCTCCCTTCCATACCCATCTCCCAAGGCTCCCAAGTGA 900
 QY 901 GAAGTCTCCAGAAAGGCTGAAACATATGATGACACAGGACCCAGACCAACAGTCAATG 960
 DB 901 GAAGTCTCCAGAAAGGCTGAAACATATGATGACACAGGACCCAGACCAACAGTCAATG 960
 QY 961 GCGGAGGTGAGAGGCATCGGTCTCGATCGGGGAGCGAGCCGAGTCCCGACGAG 1020
 DB 961 GCGGAGGTGAGAGGCATCGGTCTCGATCGGGGAGCGAGCCGAGTCCCGACGAG 1020
 QY 1021 GAAGCAAGACAGCCGCTACAGATCTGATTATGACGAGGAGAACACCATCTCTCGCAC 1080
 DB 1021 GAAGCAAGACAGCCGCTACAGATCTGATTATGACGAGGAGAACACCATCTCTCGCAC 1080
 QY 1081 GCAGCTTACGAAACGAGCAGAGACGAGAACGCGGAGAGACACAGGATCGAGACCAACGAA 1140
 DB 1081 GCAGCTTACGAAACGAGCAGAGACGAGAACGCGGAGAGACACAGGATCGAGACCAACGAA 1140
 QY 1141 GATCACCATCTCTGAAAGGCTCTTCAAAAAGAGTATAGAGATCTGGAAGAGTACG 1200
 DB 1141 GATCACCATCTCTGAAAGGCTCTTCAAAAAGAGTATAGAGATCTGGAAGAGTACG 1200
 QY 1201 GTTTATCGGTCTCTGAAACCTGCTGGATGCAACACAGAAATTACCTGGGAGATTATTA 1260
 DB 1201 GTTTATCGGTCTCTGAAACCTGCTGGATGCAACACAGAAATTACCTGGGAGATTATTA 1260
 QY 1261 AAAATAACAGATTCTTTGGGCCCCACCTTGGAGATTGTGAATCATCGCTCCCCAAGTAGG 1320
 DB 1261 AAAATAACAGATTCTTTGGGCCCCACCTTGGAGATTGTGAATCATCGCTCCCCAAGTAGG 1320
 QY 1321 AGAAGAGAGACTCGTTGGGAGGAGAGAAAAGACCGTTGGAGTGAACACAGAGTCTTG 1380
 DB 1321 AGAAGAGAGACTCGTTGGGAGGAGAGAAAAGACCGTTGGAGTGAACACAGAGTCTTG 1380
 QY 1381 GCAAGACAGAACTATACCTCAATCAAGAAAAAGAGCCCGAGGAGACCATGCTCTGACA 1440
 DB 1381 GCAAGACAGAACTATACCTCAATCAAGAAAAAGAGCCCGAGGAGACCATGCTCTGACA 1440

1441 AGAATGAGGAGGAAGAAGAACTTCTTTAAGCCTGTGTGGATTGGATGCACTCATTTGAG 1500
1441 AGAATGAGGAGGAAGAAGAACTTCTTTAAGCCTGTGTGGATTGGATGCACTCATTTGAG 1500
1501 AAAAATCTACTACTCCAGTGAAGCCCATGGATCAGGTGGGAGATTCTAAGTGGTGGAAAGCA 1560
1501 AAAAATCTACTACTCCAGTGAAGCCCATGGATCAGGTGGGAGATTCTAAGTGGTGGAAAGCA 1560
1561 GTAGGCTTCGTGTGATCTTATATGACAAATTTGAGAGGAGTTGGGAGCAGGCAAGAAAAGG 1620
1561 GTAGGCTTCGTGTGATCTTATGACAAATTTGAGAGGAGTTGGGAGCAGGCAAGAAAAGG 1620
1621 CCAAGCTCTCGTGGCCCTCGGTGGAAACCTCCAAAGACGAAGCTCGATGAAGATTATTAGAGA 1680
1621 CCAAGCTCTCGTGGCCCTCGGTGGAAACCTCCAAAGACGAAGCTCGATGAAGATTATTAGAGA 1680
1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1740
1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACCTGTTCTAGCAGCTCAGACT 1740
1741 CTGAACTTTTGGAGCTTTATGAGAAATCAAAAGGCAAAAAGGCCCAACCTGACCGCACTTC 1800
1741 CTGAACTTTTGGAGCTTTATGAGAAATCAAAAGGCAAAAAGGCCCAACCTGACCGCACTTC 1800
1801 ATGATGAACTTTGGTACAAAGATCCAGGCCAGATGAATGATGACCACTCTGCAAAATGCA 1860
1801 ATGATGAACTTTGGTACAAAGATCCAGGCCAGATGAATGATGACCACTCTGCAAAATGCA 1860
1861 GGGCAAGCAAGACGACAGCAAGAAATTAAGGCACAGCAATTTATCTGGAGAGAGGCCATCA 1920
1861 GGGCAAGCAAGACGACAGCAAGAAATTAAGGCACAGCAATTTATCTGGAGAGAGGCCATCA 1920
1921 AGCCCTGTCTGATGACAAATGCTGGCAGACTTTTTCACATACCGGATCAGAGTCT 1980
1921 AGCCCTGTCTGATGACAAATGCTGGCAGACTTTTTCACATACCGGATCAGAGTCT 1980
1981 CCCCCTGACGAACTTTTAACTGACAGGCCAACTGTTATAGAAACGATGATCAGAGT 2040
1981 CCCCCTGACGAACTTTTAACTGACAGGCCAACTGTTATAGAAACGATGATCAGAGT 2040
2041 ATATCTTTGAAAGGATTTTCTATGTTTGACATGCCCCCTGACCAATATTCACATCTGTGA 2100
2041 ATATCTTTGAAAGGATTTTCTATGTTTGACATGCCCCCTGACCAATATTCACATCTGTGA 2100
2101 AAGTAATATGATTCACATAGACTACAGATTCATTTTCAATTTGAGAGATGATGCCGGAGA 2160
2101 AAGTAATATGATTCACATAGACTACAGATTCATTTTCAATTTGAGAGATGATGCCGGAGA 2160
2161 ATTTTGTGTGAAGGGCTTGAACTCTTTTCACTGTTCTCTATTCAGAGATATTTTGGAAAT 2220
2161 ATTTTGTGTGAAGGGCTTGAACTCTTTTCACTGTTCTCTATTCAGAGATATTTTGGAAAT 2220
2221 TATGACTGGAATCTTAAAGGTCCTTTGTTGAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
2221 TATGACTGGAATCTTAAAGGTCCTTTGTTGAAGACAGCCCTCCCTGCTGCCCAAGAT 2280
2281 TTCAATTTGATGCCAGCTTTTGAAGATTTCTTCAGATGGAGAAAGGAAGTGTCTGTCCA 2340
2281 TTCAATTTGATGCCAGCTTTTGAAGATTTCTTCAGATGGAGAAAGGAAGTGTCTGTCCA 2340
2341 TGCACAGATTTCTCTGTACTTTTAAAGTGAGGACAAAGCCCTGCTGAGGAGGAGA 2400
2341 TGCACAGATTTCTCTGTACTTTTAAAGTGAGGACAAAGCCCTGCTGAGGAGGAGA 2400
2401 TTGCCAATATGCTTCACTGAGTGGAGAGCTGGAGTGGCAGAAATATGACAGAAATGCAAG 2460
2401 TTGCCAATATGCTTCACTGAGTGGAGAGCTGGAGTGGCAGAAATATGACAGAAATGCAAG 2460
2461 GCATGATTTGTTACCAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520
2461 GCATGATTTGTTACCAACCCCTGGGACGAAACCAAGCTCTGTCCGTATCGATCAACTGGATC 2520

2521 GTGAACAGTTCAACCCCGATGTGATTACTTTTCCGATTATCGTCCACTTTGGGATACGCC 2580
2521 GTGAACAGTTCAACCCCGATGTGATTACTTTTCCGATTATCGTCCACTTTGGGATACGCC 2580
2581 CTGCAAGTTGATGATTATGAGGAGACCCACAGTACAAAACTGTGGAAAGATTATGTGA 2640
2581 CTGCAAGTTGATGATTATGAGGAGACCCACAGTACAAAACTGTGGAAAGATTATGTGA 2640
2641 AACTTTCGCACCTCTAGCAATAGTCCCAAACTCAAACTGCAAACTGCAAACTGCAAACTGCA 2700
2641 AACTTTCGCACCTCTAGCAATAGTCCCAAACTCAAACTGCAAACTGCAAACTGCAAACTGCA 2700
2701 CACAGAGGAGGAGGAGCCCTCCAAAATAACGCGCAGAGAAATACAATGACGACGAGAGTAA 2760
2701 CACAGAGGAGGAGGAGCCCTCCAAAATAACGCGCAGAGAAATACAATGACGACGAGAGTAA 2760
2761 CCGTGGAGTAAAGTAGCAGCAAGATTCTGGAATACTGGCAATCCGTTCTGATGTCTGTGACG 2820
2761 CCGTGGAGTAAAGTAGCAGCAAGATTCTGGAATACTGGCAATCCGTTCTGATGTCTGTGACG 2820
2821 ATGCAATGATGCTACTGTTCTGACCCATCATATCCGCTACCACTGCTTAATGCAAT 2880
2821 ATGCAATGATGCTACTGTTCTGACCCATCATATCCGCTACCACTGCTTAATGCAAT 2880
2881 TGGCAAGTTGATAGGATATCTTTCCAAAGTCTGTTCTGTTGCAAGTGGCCATGACTC 2940
2881 TGGCAAGTTGATAGGATATCTTTCCAAAGTCTGTTCTGTTGCAAGTGGCCATGACTC 2940
2941 ATCCAAGTCAATTTAAATTTTGGAAATGAAATCTGATCATGCGCAAGAAATCATTTATCTTA 3000
2941 ATCCAAGTCAATTTAAATTTTGGAAATGAAATCTGATCATGCGCAAGAAATCATTTATCTTA 3000
3001 ACTGTGGAAATCGGAGCCCAATACGGAGACAGAAAAGTTTCAATCATGCAATCGCGGA 3060
3001 ACTGTGGAAATCGGAGCCCAATACGGAGACAGAAAAGTTTCAATCATGCAATCGCGGA 3060
3061 AGAAGGGGATTAACCTTTGATAAAATATCATGTCAAGCTTGGCCCAAGATGACCAACTC 3120
3061 AGAAGGGGATTAACCTTTGATAAAATATCATGTCAAGCTTGGCCCAAGATGACCAACTC 3120
3121 CTGAGGAGATTAACCAATGAAACGGTTGGAAATTTCTGGGTGATGCTGTTGTTGAATTTTC 3180
3121 CTGAGGAGATTAACCAATGAAACGGTTGGAAATTTCTGGGTGATGCTGTTGTTGAATTTTC 3180
3181 TGCACAGCTGCTTGTACTATTTGTTTCTAGTCTGGAAGAGGAGATTAGCAACT 3240
3181 TGCACAGCTGCTTGTACTATTTGTTTCTAGTCTGGAAGAGGAGATTAGCAACT 3240
3241 ATCGGACTGCCATTTGTCAGAAATCAGCACCTTTGCCATGCTAGCAAAAGAACTTGAACCTGG 3300
3241 ATCGGACTGCCATTTGTCAGAAATCAGCACCTTTGCCATGCTAGCAAAAGAACTTGAACCTGG 3300
3301 ATCCATTTATGCTGATGCTACCGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
3301 ATCCATTTATGCTGATGCTACCGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
3361 CAATGGCCAAATTTTGGAGCGTTAAATAGGAGCTGTTTACTTGGAGGAGGAGCTTCGAGG 3420
3361 CAATGGCCAAATTTTGGAGCGTTAAATAGGAGCTGTTTACTTGGAGGAGGAGCTTCGAGG 3420
3421 AAGCCAAAGCAGTTATTTGGAGCGCTTCTCTTTAAATGATCCGACCTTCGCGAAGTCTGGC 3480
3421 AAGCCAAAGCAGTTATTTGGAGCGCTTCTCTTTAAATGATCCGACCTTCGCGAAGTCTGGC 3480
3481 TCAATTTATCTCTCCACCCACTCAAGAGCAAAATCTGATCGCAAACTTTATTTG 3540
3481 TCAATTTATCTCTCCACCCACTCAAGAGCAAAATCTGATCGCAAACTTTATTTG 3540
3541 AACTTCTCCAGTTCTTACAAAACCTTACTGAGTTTGAAGAGCAATTTGAGTAATTTTGA 3600
3541 AACTTCTCCAGTTCTTACAAAACCTTACTGAGTTTGAAGAGCAATTTGAGTAATTTTGA 3600
3601 CTCAATGTTCCGACTTCTGGCAAGGGCATTTCAATTTAGAACTGTGGGATTTTAACTCATGTA 3660


```

3601 CTATGTTTCGACTTCTGGCAAGGCAATTCACATTTGAGAACTGTGGATTTAAACCATCTGA 3660
3661 CCCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGCTCCATAAATGCAACTGGTAGCCA 3720
3661 CCCTAGGCCCAATCAGAGAAATGGAATTCCTAGGTGCTCCATAAATGCAACTGGTAGCCA 3720
3721 CAGAGTACTTATTCATTCATTTCCAGATCATCATGAGGACACCTAACTTTGTCGGAA 3780
3721 CAGAGTACTTATTCATTCATTTCCAGATCATCATGAGGACACCTAACTTTGTCGGAA 3780
3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATCAGAGGT 3840
3781 GCTCTTTGGTGAATAATAGAACTCAGGCCAAGGTAGCGGAGGAGCTGGGCATCAGAGGT 3840
3841 ACCCCATAACCAACGACAAGACCAAGAGGCTGTGGCGCTTCGCAACAAGACCTTGGCGG 3900
3841 ACCCCATAACCAACGACAAGACCAAGAGGCTGTGGCGCTTCGCAACAAGACCTTGGCGG 3900
3901 ACCTTTTGGAAATCATTTATTTGACGCGCTGACACTGATAGGATTTGGAAATATGTTTCA 3960
3901 ACCTTTTGGAAATCATTTATTTGACGCGCTGACACTGATAGGATTTGGAAATATGTTTCA 3960
3961 CTTTCATGAATGCTGCTCTTTCCAGATTTGAAAGAAATTCATTTTGAATCAGATTTGGA 4020
3961 CTTTCATGAATGCTGCTCTTTCCAGATTTGAAAGAAATTCATTTTGAATCAGATTTGGA 4020
4021 ATGACCCCAATCCCAAGCTTCAGCAGTGTGCTTGACACTTAGAGCAGAGGAAAGAGC 4080
4021 ATGACCCCAATCCCAAGCTTCAGCAGTGTGCTTGACACTTAGAGCAGAGGAAAGAGC 4080
4081 CAGACATTCCTCTGTACAGACTCTGACAGAGTGGGCCCATCCATGCCGAACTTACA 4140
4081 CAGACATTCCTCTGTACAGACTCTGACAGAGTGGGCCCATCCATGCCGAACTTACA 4140
4141 CTGTGCTGTGTTATTTCAAGGAGAAAGAAATAGGCTGTGGAAAGACCAAGTATTTCAGC 4200
4141 CTGTGCTGTGTTATTTCAAGGAGAAAGAAATAGGCTGTGGAAAGACCAAGTATTTCAGC 4200
4201 AAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAAATAATATATTTTCCCGAGTGCC 4260
4201 AAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAAATAATATATTTTCCCGAGTGCC 4260
4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTGTAAAGAAATGAGTGGGAA 4320
4261 ATCAGAAGCGGTTTCATCGAACGGAAGTACAGACAGAGTGTAAAGAAATGAGTGGGAA 4320
4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGACACATCAAGAAATTAAGAGGGGCAT 4380
4321 GAGAGCATCAAGAGAGAGCCAGATGAGACTGACACATCAAGAAATTAAGAGGGGCAT 4380
4381 GCAAGTGTGGAGTATTTACTTGTCTAGTAACTGTGACTGTGCTATTGAGACCTAGCCT 4440
4381 GCAAGTGTGGAGTATTTACTTGTCTAGTAACTGTGACTGTGCTATTGAGACCTAGCCT 4440
4441 AGTTTTCCTGACAGCAATGAAACGAAAGTGTCTCATTTGAAATAAATAACAGAGTCAA 4500
4441 AGTTTTCCTGACAGCAATGAAACGAAAGTGTCTCATTTGAAATAAATAACAGAGTCAA 4500
4501 CTATTGTTCTTTTAATGATCTGTTTTAGTGGATGCTCTTTATTTAACAAGTATTAGATT 4560
4501 CTATTGTTCTTTTAATGATCTGTTTTAGTGGATGCTCTTTATTTAACAAGTATTAGATT 4560
4561 TTTCTTCTATTATTAACGAAATCTGATTTGTTGGAATGTGCTTCTTTTATTTTGG 4620
4561 TTTCTTCTATTATTAACGAAATCTGATTTGTTGGAATGTGCTTCTTTTATTTTGG 4620
4621 CTCCTTTAAATAATAAATCAAGAGCATATTTCTATGTGGAATAGATCTGTTTTTCCAT 4680
4621 CTCCTTTAAATAATAAATCAAGAGCATATTTCTATGTGGAATAGATCTGTTTTTCCAT 4680
4681 CTGTGTCCAGATTTGACCTAGACTTTCATTTGACAGATTAATAAATTTGACTTTACTAG 4740
```

```

4681 CTGTGTCCAGATTTGTGACCTAGACTTTTCAATTGACAAGTAAATAAATTGACTTTACTAG 4740
4741 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4764
4741 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4764

RESULT 5
HSM807870      4650 bp mRNA linear PRI 30-AUG-2003
LOCUS          Homo sapiens mRNA; cDNA DKF2p686F16262 (from clone DKF2p686F16262).
DEFINITION    BX647724
ACCESSION     BX647724.1 GI:34366881
VERSION
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE     1 (bases 1 to 4650)
AUTHORS      Foustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM      The German Human cDNA Consortium
TITLE         Direct Submission
JOURNAL       Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
COMMENT       Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKF2p686F16262) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
             Location/Qualifiers
             1..4650
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="DKF2p686F16262"
             /cissue_type="human cerebellum"
             /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
             DH10B; sites SfiIA + SfiIB"
             /dev_stage="adult"
             polyA_signal 4613..4618
             polyA_site 4629
             ORIGIN

Query Match          91.2%; Score 4343.4; DB 8; Length 4650;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 4471; Conservative 0; Mismatches 11; Indels 111; Gaps 1;

QY 58 GAGAGCTTTTATAGTCTCTTTCCCGGGATGTGAAGATACAGAAATGACTGTGAAT 117
DB 151 GAGTACTTTTCCAGTTCCTTTCCCGGGATGTGAAGATACAGAAATGACTGTGAAT 210
QY 118 CAACCATATCATCAGGAGCTGTAATCTAGTGGAGAGTGTAGCGTGTGCATCTTCA 177
DB 211 CAACCATATCATCAGGAGCTCTCTGAGCTTATTTCTCTGTGGAGAGTGTAGCGTGTGCATCTTCA 270
QY 178 CTATGATATGAGGAGCTCTCTGAGCTTATTTCTCTGTGGAGAGTGTAGCGTGTGCATCTTCA 237
DB 271 CTATGATATGAGGAGCTCTCTGAGCTTATTTCTCTGTGGAGAGTGTAGCGTGTGCATCTTCA 330
QY 238 GGAAACATCATGATCAGGGGAAACACATGTCTACAGAAATGTCTTCCACCGGGACGAGGC 297
DB 331 GGAAACATCATGATCAGGGGAAACACATGTCTACAGAAATGTCTTCCACCGGGACGAGGC 390
QY 298 GTCCCCGAGGACGAGGAGGACATGAGCCAGACCTCAGACCATCTTTAGCCCCCAAA 357
```

Db 391 GTCCCGAGGACGAGGAGACATGGAGCCAGACCCCTCAGCACCATCTCTTTAGGCCCAAA 450
QY 358 ATCTGAGGCTGTTTACCCCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTCCAGTG 417
Db 451 ATCTGAGGCTGTTTACCCCTCAGCAGCTCTCTGTGCAATATCAATATGAACCTCCAGTG 510
QY 418 CCCCTTCCACCACTTCTCMAACTCTCCAGCCCCCAATTTCTCCCTCCACGACAGACT 477
Db 511 CCCCTTCCACCACTTCTCMAACTCTCCAGCCCCCAATTTCTCCCTCCACGACAGACT 570
QY 478 TTGTACCTTCCCCCAACCATGCTCTCCGTACGCGCAAGGCCCTCTTCCCCCTGCCCCAA 537
Db 571 TTGTACCTTCCCCCAACCATGCTCTCCGTACGCGCAAGGCCCTCTTCCCCCTGCCCCAA 630
QY 538 TCAGGCGGCTTTCGCCCAACCAACAGATGAGGCAACCCCTTCCAGTTCTCTCTGTTTTC 597
Db 631 TCAGGCGGCTTTCGCCCAACCAACAGATGAGGCAACCCCTTCCAGTTCTCTCTGTTTTC 690
QY 598 CTCCCATGCCACCAATGCTTGTCTTAATAACCCCGCAGTCCCTGGGGCACTCTCTG 657
Db 691 CTCCTATGCCACCAACCAATGCTTGTCTTAATAACCCCGCAGTCCCTGGGGCACTCTCTG 750
QY 658 GACAGGCACTTTCCTCTTCATGATGCCCTCTCCCTCCATGCTCATCCCCCGCCCCCTC 717
Db 751 GACAGGCACTTTCCTCTTCATGATGCCCTCTCCCTCCATGCTCATCCCCCGCCCCCTC 810
QY 718 CAGTCATGCCGAGAGGTTAATTATCAGTACCTCCGGGCTATTCTCACCACAACCTTCC 777
Db 811 CAGTCATGCCGAGAGGTTAATTATCAGTACCTCCGGGCTATTCTCACCACAACCTTCC 870
QY 778 CACCTCCAGTTTAAATAGTTTCCAGACAACCCCTAGTTCTTCTCGCCAGTGCTAAATA 837
Db 871 CACCTCCAGTTTAAATAGTTTCCAGACAACCCCTAGTTCTTCTCGCCAGTGCTAAATA 930
QY 838 ACAGCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAGTG 897
Db 931 ACAGCAGTAGTCTCATTTTCAGACATCTCCCTCCATACCCACTCCCAAGGCTCCCAGTG 990
QY 898 AGAAGAGTCCCGAAGAGCTGAACATATGATGACCAAGGACCGAGACCAAGTC 957
Db 991 AGAAGAGTCCCGAAGAGCTGAACATATGATGACCAAGGACCGAGATCACAGTC 1050
QY 958 ATGGCGAGTGAGAGCATCGTCTCCGTGATCGCGGAGCGAGCGCCAGTCCCGACA 1017
Db 1051 ATGGCGAGTGAGAGCATCGTCTCCGTGATCGCGGAGCGAGCGCCAGTCCCGACA 1110
QY 1018 GGAGAGACAAGACAGCGGTACAGATCTGATTATGACCGAGGGAGAACCAATCTCGCC 1077
Db 1111 GGAGAGACAAGACAGCGGTACAGATCTGATTATGACCGAGGGAGAACCAATCTCGCC 1170
QY 1078 ACCGCACTACGAACGGAGCAGAGAGCGAGAACGGGAGAGACACAGGCATCGAGACAAC 1137
Db 1171 ACCGCACTACGAACGGAGCAGAGAGCGAGAACGGGAGAGACACAGGCATCGAGACAAC 1230
QY 1138 GAGATCACCATCTCTGGAAGCTCTTACAAAAGAGATATAGAGATCTGGAAGGAT 1197
Db 1231 GAGATCACCATCTCTGGAAGGCTCTTACAAAAGAGATATAGAGATCTGGAAG----- 1285
QY 1198 ACGGTTTATCGGTGTTCTGAACTGCTGGATGCAACACAGAAATTACCTGGGGAGATTA 1257
Db 1286 ----- 1285
QY 1258 TTAATAATACAGATTTCTTGGGCCCCACCCCTGGAGATTTGGAATCATCTCCCTCCCAAGTA 1317
Db 1286 -----TCGCTCCCCCAAGTA 1299
QY 1318 GCGAGAGAGAGAGCTCTGTTGGGAGGAGAAAAGACCGTCTGAGTGAACAACAGAGTT 1377
Db 1300 GCGAGAGAGAGAGCTCTGTTGGGAGGAGAAAAGACCGTCTGAGTGAACAACAGAGTT 1359
QY 1378 CTGGCAAGACAAGAACTATACCTCAATCAAGGAAAAGAGCCCGAGGAGACCATGCTG 1437
Db 1360 CTGGCAAGACAAGAACTATACCTCAATCAAGGAAAAGAGCCCGAGGAGACCATGCTG 1419

QY 1438 ACAAGATGAGGAGGAAGAAGAAGAACTTTCTTAAGCCTGTGTGGATTTCGATGCATCTATT 1497
Db 1420 ACAGATGAGGAGGAAGAAGAAGAACTTTCTTAAGCCTGTGTGGATTTCGATGCATCTATT 1479
QY 1498 CAGAAAACTACTACTCCAGTGACCCCATGGATCAGGTGGAGATTCTCAGTGGTTGGAA 1557
Db 1480 CAGAAAACTACTACTCCAGTGACCCCATGGATCAGGTGGAGATTCTCAGTGGTTGGAA 1539
QY 1558 CGAGTAGGCTTCGTACTTATATGACAAATTTGAGAGGAGTTGGGAGACGCGCAAGAA 1617
Db 1540 CGAGTAGGCTTCGTACTTATATGACAAATTTGAGAGGAGTTGGGAGACGCGCAAGAA 1599
QY 1618 AGGCCAAAAGCTCTCGGCTCCGTGGGAACCTCCAAAAGACGAAGCTCGATGAAGATTAG 1677
Db 1600 AGGCCAAAAGCTCTCGGCTCCGTGGGAACCTCCAAAAGACGAAGCTCGATGAAGATTAG 1659
QY 1678 AGAGTTCCAGTGAATCCGAGTGTGATGTGATGAGGACAGCACCTGTTCTAGCAGTCTAG 1737
Db 1660 AGAGTTCCAGTGAATCCGAGTGTGATGTGATGAGGACAGCACCTGTTCTAGCAGTCTAG 1719
QY 1738 ACTCTGAAGTTTTCGACGTTATTCAGAAATCAAAACGCAAAAGGCCCAACCTGACCGAC 1797
Db 1720 ACTCTGAAGTTTTCGACGTTATTCAGAAATCAAAACGCAAAAGGCCCAACCTGACCGAC 1779
QY 1798 TTTCATGATGAACCTTTGGTACAAACGATCCAGGCGCAGATGAATGATGACCACTCTGCAAT 1857
Db 1780 TTTCATGATGAACCTTTGGTACAAACGATCCAGGCGCAGATGAATGATGACCACTCTGCAAT 1839
QY 1858 GCAGCGCAAGCAAGACGACACAGGAATPAGGCAAGCAATTTATCTGTGAGAGAGGCCA 1917
Db 1840 GCAGCGCAAGCAAGACGACACAGGAATPAGGCAAGCAATTTATCTGTGAGAGAGGCCA 1899
QY 1918 TCAGGCCCTGCTCTTATGACCAACATGCTGGCAGACTTTTCCACTACCGGATCACAG 1977
Db 1900 TCAAGCCCTGCTCTTATGACCAACATGCTGGCAGACTTTTTCACACTACCGGATCACAG 1959
QY 1978 TCTCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAATACGATGATCAG 2037
Db 1960 TCTCCCGCTACGAACTTTTAACTGACAGGCCAACTGTTATAGAATACGATGATCAG 2019
QY 2038 AGTATATCTTTGAAGATTTTCTATGTTTGCACATGCCCCCTGACCAATATTCCTGT 2097
Db 2020 AGTATATCTTTGAAGATTTTCTATGTTTGCACATGCCCCCTGACCAATATTCCTGT 2079
QY 2098 GTAAAGTAATTAGATTCAACATAGACTACAGATTCAATTTCAATGAGAGATGATCCGG 2157
Db 2080 GTAAAGTAATTAGATTCAACATAGACTACAGATTCAATTTCAATGAGAGATGATCCGG 2139
QY 2158 AGAATTTTGTGAAAGGCTTGAACCTCTTTTCACTGTTCTCTTATTCAGAGATATTGG 2217
Db 2140 AGAATTTTGTGAAAGGCTTGAACCTCTTTCACTGTTCTCTTATTCAGAGATATTGG 2199
QY 2218 AATTATATGACTGGAATCTTAAAGGCTCTTTGTTGAAGACAGCCCTCCCTGTCGCCAA 2277
Db 2200 AATTATATGACTGGAATCTTAAAGGCTCTTTGTTGAAGACAGCCCTCCCTGTCGCCAA 2259
QY 2278 GATTTCAATTTCAAGCAGCTTTTGAAGATTCTTCCAGATGGAGAAAGGAGTCTGT 2337
Db 2260 GATTTCAATTTCAAGCAGCTTTTGAAGATTCTTCCAGATGGAGAAAGGAGTCTGT 2319
QY 2338 CCATGACACAGATTCTCTGTACTTGTAAAGGTGACGCAAAAGCCCTGGTGCCTGAGGAG 2397
Db 2320 CCATGACACAGATTCTCTGTACTTGTAAAGGTGACGCAAAAGCCCTGGTGCCTGAGGAG 2379
QY 2398 AGATTGCCAATATGCTTCAAGTGGAGAGCTGGAGTGGCAGAAATATGCGAAGAAATGCA 2457
Db 2380 AGATTGCCAATATGCTTCAAGTGGAGAGCTGGAGTGGCAGAAATATGCGAAGAAATGCA 2439
QY 2458 AAGGCATGATTGTATCAACCCCTGGGACGAAACCAAGCTCTGTCCGCTATCATCAACTGG 2517
Db 2440 AAGGCATGATTGTATCAACCCCTGGGACGAAACCAAGCTCTGTCCGCTATCATCAACTGG 2499

QY 2518 ATCGTGAACAGTTCAACCCCGATGTGATTACTTTTCGGATTATCGTCCACCTTTGGGATAC 2577
Db 2500 ATCGTGAACAGTTCAACCCCGATGTGATTACTTTTCGGATTATCGTCCACCTTTGGGATAC 2559
QY 2578 GCCCTGCACAGTTGAGTTATGCAAGGAGACCACAGTACCAAAACCTGTGGAAGAGTTATG 2637
Db 2560 GCCCTGCACAGTTGAGTTATGCAAGGAGACCACAGTACCAAAACCTGTGGAAGAGTTATG 2619
QY 2638 TGAACCTTGCACCTCTAGCAAAATAGTCCCAAAAGTCAAACTGACAAACAGAAC 2697
Db 2620 TGAACCTTGCACCTCTAGCAAAATAGTCCCAAAAGTCAAACTGACAAACAGAAC 2679
QY 2698 TGSCACAGAGGAGGAGGAGCCCTCCAAAAATACGSCAGAGAAATCAATGAGACGAGAAG 2757
Db 2680 TGSCACAGAGGAGGAGGAGCCCTCCAAAAATACGSCAGAGAAATCAATGAGACGAGAAG 2739
QY 2758 TAAAGGTGGAGCTAAGTAGCAAGGATTTCTGGAAAACTGSCATCCGTTCTGATGTCGTC 2817
Db 2740 TAAAGGTGGAGCTAAGTAGCAAGGATTTCTGGAAAACTGSCATCCGTTCTGATGTCGTC 2799
QY 2818 AGCATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCTTAATGC 2877
Db 2800 AGCATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCTTAATGC 2859
QY 2878 ATTTGACAAAGTTGATAGGATATCTTCCAAAGATCGTTGTCGTCGAGCTGGCCATCA 2937
Db 2860 ATTTGACAAAGTTGATAGGATATCTTCCAAAGATCGTTGTCGTCGAGCTGGCCATCA 2919
QY 2938 CTCATCCAAAGTCATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGCAGGAATTCATTAAT 2997
Db 2920 CTCATCCAAAGTCATCAATTTAAATTTTGGAAATGAATCCTGATCATGCGCAGGAATTCATTAAT 2979
QY 2998 CTAACTGTGGAATTCGGGAGCCCAATACGGAGACAGAAAGTTTCATGACATGC 3057
Db 2980 CTAACTGTGGAATTCGGGAGCCCAATACGGAGACAGAAAGTTTCATGACATGC 3039
QY 3058 GGAAGAAAGGGAATTAACCTTGATATAATATCATGTGTCAGCCCTTGGCCAGAGATGACCCAA 3117
Db 3040 GGAAGAAAGGGAATTAACCTTGATATAATATCATGTGTCAGCCCTTGGCCAGAGATGACCCAA 3099
QY 3118 CTCCTTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTGTTGAAT 3177
Db 3100 CTCCTTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGGTGATGCTGTTGTTGAAT 3159
QY 3178 TTCTGACAGCGTCCATTTGTACTATTTGTTCTAGTCTGGAAGAAAGAGGANTTAGCAA 3237
Db 3160 TTCTGACAGCGTCCATTTGTACTATTTGTTCTAGTCTGGAAGAAAGAGGANTTAGCAA 3219
QY 3238 CCTATCGGATGCGCATTTGTCAGAAATCAGCACCTTGCCATGCTAGCAAAAGAACTTGAAC 3297
Db 3220 CCTATCGGATGCGCATTTGTCAGAAATCAGCACCTTGCCATGCTAGCAAAAGAACTTGAAC 3279
QY 3298 TGGATCCATTTATGCTGTATGCTACGGGCTGACCTTTGTAGAGAAATCGGACCTTCGAC 3357
Db 3280 TGGATCGATTTATGCTGTATGCTACGGGCTGACCTTTGTAGAGAAATCGGACCTTCGAC 3339
QY 3358 ATGCAATGCGCAATGCTTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGCCTGG 3417
Db 3340 ATGCAATGCGCAATGCTTTTGAAGCGTTAATAGGAGCTGTTTACTTTGGAGGGAAGCCTGG 3399
QY 3418 AGGAGCCAAAGCAGTTATTTGAGCGCTTCTCTTTAATGATCGGACCTGCGGGAAGTCT 3477
Db 3400 AGGAGCCAAAGCAGTTATTTGAGCGCTTCTCTTTAATGATCGGACCTGCGGGAAGTCT 3459
QY 3478 GGCCTCAATTTATCTCTCCACCCACTCCAACTACAAGAGCAAAATCTGATCGACAACTTA 3537
Db 3460 GGCCTCAATTTATCTCTCCACCCACTCCAACTACAAGAGCAAAATCTGATCGACAACTTA 3519
QY 3538 TTGAAACTCTTCCAGTTCTACAAAACTTACTGAGTTTGAAGAGCAATTTGAGGTAATTT 3597
Db 3520 TTGAAACTCTTCCAGTTCTACAAAACTTACTGAGTTTGAAGAGCAATTTGAGGTAATTT 3579
QY 3598 TTACTCATGTTGCGACTTCTTGGCAAGGGCATTTTCATTTGAGAACTGTGGGATTTAACCATC 3657

Db 3580 TTACTCATGTTGCGACTTCTTGGCAAGGCATTCACATTTGAGAACTGTGGGATTTAACCATC 3639
QY 3658 TGAACCTTGGCCCACAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAG 3717
Db 3640 TGAACCTTGGCCCACAATCAGAGAAATGGAATTCCTAGGTGACTCCATAATGCAACTGGTAG 3699
QY 3718 CCACAGAGTACTTATTCATTTCCAGAGTCAATCATGAGGACACTTAACTTTGTTGC 3777
Db 3700 CCACAGAGTACTTATTCATTTCCAGATCATCATGAGGACACTTAACTTTGTTGC 3759
QY 3778 GAACTCTTTTGGTGAATAATAGAACTCAGGCCAAAGGTAGCGAGGAGCTGGGCATGCAGG 3837
Db 3760 GAACTCTTTTGGTGAATAATAGAACTCAGGCCAAAGGTAGCGAGGAGCTGGGCATGCAGG 3819
QY 3838 AGTACGCCATAACCAACGACAGACCAAGAGGCTGTGGCGTTTCGACCAAGACCTTTGG 3897
Db 3820 AGTACGCCATAACCAACGACAGACCAAGAGGCTGTGGCGTTTCGACCAAGACCTTTGG 3879
QY 3898 CGGACCTTTTGGANCAATTTATTTGAGCGCTGTACACTGATAGGATTTGGAAATATGTTT 3957
Db 3880 CGGACCTTTTGGANCAATTTATTTGAGCGCTGTACACTGATAGGATTTGGAAATATGTTT 3939
QY 3958 ATACTTTTCATGAATGTCTCTTTTCCACGATTTGAAAGAAATTCATTTTGAATCAGGATT 4017
Db 3940 ATACTTTTCATGAATGTCTCTTTTCCACGATTTGAAAGAGTTTCATTTTGAATCAGGATT 3999
QY 4018 GGAATGACCCCAATCCCAAGCTTTCAGCAGTGTGTGTTGACACTTAGGACAGAGGAAAG 4077
Db 4000 GGAATGACCCCAATCCCAAGCTTTCAGCAGTGTGTGTTGACACTTAGGACAGAGGAAAG 4059
QY 4078 AGCCAGACATTTCTCTGTACAAGACTCTTCAGACAGTGGGCCCATCCCATGCCGAACCT 4137
Db 4060 AGCCAGACATTTCTCTGTACAAGACTCTTCAGACAGTGGGCCCATCCCATGCCGAACCT 4119
QY 4138 ACACTGTGCTGTGTTATTTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGCAACAAATATTC 4197
Db 4120 ACACTGTGCTGTGTTATTTTCAAGGGAGAAAGAAATAGGCTGTGGGAAAGCAACAAATATTC 4179
QY 4198 AGCAAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAAAAATATAATTTTCCCAGATGG 4257
Db 4180 AGCAAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAAAAATATAATTTTCCCAGATGG 4239
QY 4258 CCATCAGAAAGCGGTTTCATCGAAACGGAGTACAGACAGAGTTTAAAGAAATGAGGTGG 4317
Db 4240 CCATCAGAAAGCGGTTTCATCGAAACGGAGTACAGACAGAGTTTAAAGAAATGAGGTGG 4299
QY 4318 AAAGAGACATCAAGAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGGAGGG 4377
Db 4300 AAAGAGACATCAAGAGAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGGAGGG 4359
QY 4378 CATGCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGTGCTTGTCTATTGAGACCTAG 4437
Db 4360 CATGCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGTGCTTGTCTATTGAGACCTAG 4419
QY 4438 CCTAGTTTTCCTGCGACAGAAATGAAACGAGTGTGCTCATTTGAAATAAATAACAGAGTCAA 4497
Db 4420 CCTAGTTTTCCTGCGACAGAAATGAAACGAGTGTGCTCATTTGAAATAAATAACAGAGTCAA 4479
QY 4498 TCGCTATTGTTGCTTTTAAATGATCTGTTTTTGTAGCTGGTCTTTTATTACAAAGTATTAG 4557
Db 4480 TCGCTATTGTTGCTTTTAAATGATCTGTTTTTGTAGCTGGTCTTTTATTACAAAGTATTAG 4539
QY 4558 ATTTTTTCTTCTAATTTAAACGGAACACTTGAATTTTGGTGAATGTGCAATCTCTCTTTATT 4617
Db 4540 ATTTTTTCTTCTAATTTAAACGGAACACTTGAATTTTGGTGAATGTGCAATCTCTCTTTATT 4599
QY 4618 TTGCTCTTTTAAATAATAATAATTTCAAGAGCATATA 4650
Db 4600 TTGCTCTTTTAAATAATAATAATTTCAAGAGCATATA 4632

BC054003	Homo sapiens	4069 bp	mrna	linear	PRI 20-MAY-2005
LOCUS	IMAGE:6148714, complete cds.				
DEFINITION	Homo sapiens ribonuclease III, nuclear, mRNA (cdna clone MGC:61568				
ACCESSION	BC054003				
VERSION	BC054003.1 GI:32450505				
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 4069)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shervenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Greenwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
CONSRMTM	Mammalian Gene Collection Program Team				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 4069)				
AUTHORS	NIH MGC Project				
CONSRMTM	Direct Submission				
TITLE	Submitted (24-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA				
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov				
REMARK	Contact: MGC help desk				
COMMENT	Email: cgaps-femail.nih.gov Tissue Procurement: ATCC/DCTD/DTF cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Lari,P., Legaspi,R., Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.liml.gov Series: IRAK Plate: 115 Row: 1 Column: 8 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21359821.				
source	1. .4069 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:61568 IMAGE:6148714"				
gene	/tissue_type="Skin, melanotic melanoma." /clone_lib="NIH MGC_72" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1. .4069 /gene="RNASEN" /note="synonyms: HSA242976, RN3, DROSHA, RNASE3L, Etch12" /db_xref="GeneID:29102" /db_xref="MIM:608828" 248. .3826 /gene="RNASEN" /codon_start=1 /product="RNASEN protein" /protein_id="AAH54003.1" /db_xref="GI:32450506" /db_xref="GeneID:29102" /db_xref="MIM:608828" /translation="MMQGTCHMSFHRGCGPRGGRGHGARPSPFRPQNLRLHP QQPVQYQEPSPAPSTTFSNPAENFLPRDFVFPFPPMPSPAGPLPPCIRPFP NPMWRHPFVPVPCFPMPPMPCPNPNPVPAGAPGQGTFFPMMPSPGPHPPPPVM PQOVNYPGYSHENPFPSPNSFONNPPSFLPSANSSSPHRLHLPLPYLPKPAPSE RSPRLKHYDDHRRHDSHGGERHSIDREERGESDPRRODSRYSDYDRGTSPS RHRYSRERERHRHRDNRSPSLERSYKEYKRGSGRSPSREKGRARWEEKDR WSDNOSGDKNNTYSIKKEPEETMPDKNEEEELLPVWIKTTHSENYSYSDPMDQ VGDSTVGTLSRLDLYDFEELSGRQEKAKAAPPEPPTKLDLEDLESSSECS DEBDSYSSSSEVEFDVIAEIKRKAHPDRLDHLYNDPQMDGDLCKSAKARRT GIRHSIYGCERAIKPCRPMTNAGRLFVRIIVTSPPTFLTRPTVIEYDDHEYIPEG FSMFAHALTNIPLCVIRFNIDYTHFIEEMMPENFVKGLELFLRDLRLDLYD ANMLKPLEDSPPCKQYAECKQYATVNPPTKPSNVIQDLDREQNFDPDVIITPPIVHGI REVTVELSSQGWKTSYVKLRHLLANSVKQKTDKQKLAQREELQKIKQKNTMR CLAMTHPSHLNFMNPDHARNSLNCGRPKQPKYDRKVHMHMKKGTINTLINMSR REVTVELSSQGWKTSYVKLRHLLANSVKQKTDKQKLAQREELQKIKQKNTMR LQDDPTPSRINHERLSEPLGDAVVEFLTSLVLYLPSLESGSLATVYATVQNOHL AMLAKKLELDLRFMLYAHGPDLCRESDLRHMANCFEALIGVYLSLEAKQLFQRL LFNDPDLREVWLNLYFLHPLQEPNTDRLIETSPVLQKLTFEFSAIGVIFTHVRLLA RAFTLRVTGFNHLTLGHQRMFLGDSIMQLVATYELFTHFPDHHGHLLTLRRSLVN NRTQAKVAELGMQVWSYLLNSNCCCLLRPSLFLQTMNVCVSLK"				
CDS	ORIGIN				
	Query Match	75.6%	Score 3600;	DB 8;	Length 4069;
	Best Local Similarity	97.0%	Pred. No. 0;		
	Matches 3724;	Conservative	0;	Mismatches	5; Indels 111; Gaps 1;
QY	1	CTGCTCTTGCTAGCTCGGTAGTAGCTGCTGCTCTGCTGACGCGGATCTCGGGCCCGGAG	60		
Db	3	CTGCTCTTGCTAGCTCGGTAGTAGCTGCTGCTGACGCGGATCTCGGGCCCGGAG	62		
QY	61	AGCCTTTTATAGGTGTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	120		
Db	63	AGCCTTTTATAGGTGTGCTTTTCCCGGGGATGTGAAGGATACAGAAATGACTGTGAATCAA	122		
QY	121	CCCATATCATCAAGGAGCTGATAATCTAGTGAAGATTAGAGCTGTGCATCTTCACTA	180		
Db	123	CCCATATCATCAAGGAGCTGATAATCTAGTGAAGATTAGAGCTGTGCATCTTCACTA	182		
QY	181	TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGCATATCCAGCGGA	240		
Db	183	TGATATGAGGAGCTCTCTGAGCTTATATCTCTGTGGAAGATGTGCATATCCAGCGGA	242		
QY	241	ACATCATGATGAGGAGAAACATGATGTCAAGAAATGTGTTTCCACCGGACGAGGCGTC	300		
Db	243	ACATCATGATGAGGAGAAACATGATGTCAAGAAATGTGTTTCCACCGGACGAGGCGTC	302		
QY	301	CCGAGGACGAGGAGACATGAGGACGACCCCTCAGCACATCTTTAGGCCCAAAATC	360		
Db	303	CCGAGGACGAGGAGACATGAGGACGACCCCTCAGCACATCTTTAGGCCCAAAATC	362		
QY	361	TGAGGCTGCTTCCCTCAGCAGCCCTCTGTGCAATATCAATATGAACCTTCCAAGTGCCC	420		
Db	363	TGAGGCTGCTTCCCTCAGCAGCCCTCTGTGCAATATCAATATGAACCTTCCAAGTGCCC	422		
QY	421	CTTCCACCACTTTTCTCAAACTCTCCAGCCCCCAATTTTCTCCCTCCACGACGACGTTTG	480		

Db 423 |||||CTTCCACCACTTCTCAACTCTCCAGCCCGCAATTTCTCCCTCCACGACGAGCTTTC 482
Qy 481 TACCTTTCGCCCCACCCATGCTCCGTCAGCGCAAGCCCTCTTCCCTCCCTGCCCCAATCA 540
Db 483 TACCTTTCGCCCCACCCATGCTCCGTCAGCGCAAGCCCTCTTCCCTCCCTGCCCCAATCA 542
Qy 541 GCGCGCTTTCGCCCAACCCACGATGAGGACCCCTTCCAGTTCCTCTTGTGTTTCTTC 600
Db 543 GCGCGCTTTCGCCCAACCCACGATGAGGACCCCTTCCAGTTCCTCTTGTGTTTCTTC 602
Qy 601 CCATGCCACCAAAATGCTTCTTAATAACCCCGCCAGTCCCTGGGGACACCTCTGGAC 660
Db 603 CCATGCCACCAAAATGCTTCTTAATAACCCCGCCAGTCCCTGGGGACACCTCTGGAC 662
Qy 661 AAGGCACTTTCCTTCATGATGCCCTTCCATGCTCCCTCCATGCTCATCCCCCGCCCTCCAG 720
Db 663 AAGGCACTTTCCTTCATGATGCCCTTCCATGCTCCCTCCATGCTCATCCCCCGCCCTCCAG 722
Qy 721 TCATGCCGAGCAGGTAAATTATCAGTACCTCTCGGGCTATTCTCACCACAACTTCCCAC 780
Db 723 TCATGCCGAGCAGGTAAATTATCAGTACCTCTCGGGCTATTCTCACCACAACTTCCCAC 782
Qy 781 CTCCTCAGTTTAAATAGTTTCCAGAAACAACCTAGTTCTTCTTCCTGCCCGAGTCTAAATAACA 840
Db 783 CTCCTCAGTTTAAATAGTTTCCAGAAACAACCTAGTTCTTCTTCCTGCCCGAGTCTAAATAACA 842
Qy 841 GCAGTAGTCTCATTTTCAGACATCTCCCTCCATPACCCACTCCCAAAAGGCTCCAGTGAGA 900
Db 843 GCAGTAGTCTCATTTTCAGACATCTCCCTCCATPACCCACTCCCAAAAGGCTCCAGTGAGA 902
Qy 901 GAAGTCTCCAGAAAGGCTGAACAATAATGATGACCAAGGACCCAGACCAAGTCTATG 960
Db 903 GAAGTCTCCAGAAAGGCTGAACAATAATGATGACCAAGGACCCAGACCAAGTCTATG 962
Qy 961 GCGAGGTGAGAGGATCGCTCCCTCGATCGGGGAGGAGGCCGAGTCCCGACAGGA 1020
Db 963 GCGAGGTGAGAGGATCGCTCCCTCGATCGGGGAGGAGGCCGAGTCCCGACAGGA 1022
Qy 1021 GAAGACAGACAGCCGGTACAGATCTGATTTAGCCGAGGGAGAACCAATCTCGCCACC 1080
Db 1023 GAAGACAGACAGCCGGTACAGATCTGATTTAGCCGAGGGAGAACCAATCTCGCCACC 1082
Qy 1081 GCAGTACGAAACGAGCAGAGCGAGAAACGGAGAGACACAGGCAATCAGAGAACCGAA 1140
Db 1083 GCAGTACGAAACGAGCAGAGCGAGAAACGGAGAGACACAGGCAATCAGAGAACCGAA 1142
Qy 1141 GATCACCATCTCTGGAAGGCTCTACAAAAGAGTATAGAGATCTGGAAGGATTACG 1200
Db 1143 GATCACCATCTCTGGAAGGCTCTACAAAAGAGTATAGAGATCTGGAAG----- 1194
Qy 1201 GTTTATCGGTGTTCTCTGAACCTGCTGGATGCACACAGAAATTATTA 1260
Db 1195 ----- 1194
Qy 1261 AAAATACAGATTCTGGGCCCCACCCCTGGAGATTGTGAATCATCGCTCCCAAGTAGGG 1320
Db 1195 -----TCGCTCCCAAGTAGGG 1211
Qy 1321 AGAAGAGAGAGCTCTTGGGAGGAGAAAGAACCCGTTGGAGTGAACACAGAGTTCTG 1380
Db 1212 AGAAGAGAGAGCTCTTGGGAGGAGAAAGAACCCGTTGGAGTGAACACAGAGTTCTG 1271
Qy 1381 GCAAGACAAAGAACTACTCTCAATCAAGAAAAGAGCCGAGGAGACCATGCTGTGACA 1440
Db 1272 GCAAGACAAAGAACTACTCTCAATCAAGAAAAGAGCCGAGGAGACCATGCTGTGACA 1331
Qy 1441 AGAATGAGGAGGAGAAAGAAAGAACTTTAAGCCTGTGTGGATTGATGCACTCATTTAG 1500
Db 1332 AGAATGAGGAGGAGAAAGAAAGAACTTTAAGCCTGTGTGGATTGATGCACTCATTTAG 1391
Qy 1501 AAAACTACTCTCAGTGACCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAAACGA 1560
|||||

Db 1392 AAAACTACTCTCAGTGACCCCATGGATCAGGTGGGAGATTCTACAGTGGTTGGAACGA 1451
Qy 1561 GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGAGTTGGGGAGAGCAAGAAAGG 1620
Db 1452 GTAGGCTTCGTGACTTATATGACAAATTTGAGGAGAGTTGGGGAGAGCAAGAAAGG 1511
Qy 1621 CCAAAGCTCTCGGCTCGGTGGRAACCTCCAAAGACGAGCTCGATGAAGATTAGAGA 1680
Db 1512 CCAAAGCTCTCGGCTCGGTGGAAACCTCCAAAGACGAGCTCGATGAAGATTAGAGA 1571
Qy 1681 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACTCTTCTAGCAGCTCAGACT 1740
Db 1572 GTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGACAGCACTCTTCTAGCAGCTCAGACT 1631
Qy 1741 CTGAAGTTTTGAGCTTATTCAGAAATCAAAAGGAAAGGCCCACCTCGACCGACTTC 1800
Db 1632 CTGAAGTTTTGAGCTTATTCAGAAATCAAAAGGAAAGGCCCACCTCGACCGACTTC 1691
Qy 1801 ATGATGAATCTTTGGTAAACAGATCCAGGCGCAGATGAATGATGACCACTCTGCAAAATGCA 1860
Db 1692 ATGATGAATCTTTGGTAAACAGATCCAGGCGCAGATGAATGATGACCACTCTGCAAAATGCA 1751
Qy 1861 GGGCAAGGCAAGACCGCACAGGAATTAGGCAAGCAATTTATCTCTGGAGAGAGGCCATCA 1920
Db 1752 GGGCAAGGCAAGACCGCACAGGAATTAGGCAAGCAATTTATCTCTGGAGAGAGGCCATCA 1811
Qy 1921 AGCCCTGTCTCTATGACCAACATGCTGGGAGACTTTTCCACTACCGGATCAGAGTCT 1980
Db 1812 AGCCCTGTCTCTATGACCAACATGCTGGGAGACTTTTCCACTACCGGATCAGAGTCT 1871
Qy 1981 CCGCGCTACGAACTTTTAACTGACAGGCGCACTGTATTAGNATACGATCAGAGT 2040
Db 1872 CCGCGCTACGAACTTTTAACTGACAGGCGCACTGTATTAGNATACGATCAGAGT 1931
Qy 2041 ATATCTTTGAAAGGATTTCTATGTTGACATGCCCCCTCGACCAATATTTCCACTGTGTA 2100
Db 1932 ATATCTTTGAAAGATTTCTATGTTTGGACATGCCCCCTCGACCAATATTTCCACTGTGTA 1991
Qy 2101 AAGTAATAGATTCAACATAGACTACACATTTCACTGTTCTCTTTCAGAGATATTTTGGAAAT 2160
Db 1992 AAGTAATAGATTCAACATAGACTACACATTTCACTGTTCTCTTTCAGAGATATTTTGGAAAT 2051
Qy 2161 ATTTTGTGTGAAAGGGCTTGAACCTTTTCACTGTTCTCTTTCAGAGATATTTTGGAAAT 2220
Db 2052 ATTTTGTGTGAAAGGGCTTGAACCTTTTCACTGTTCTCTTTCAGAGATATTTTGGAAAT 2111
Qy 2221 TATATGACTGGAAATCTTAAAGGTCTTCTTGTGAAAGACAGCCCTCTCTGCTGCCAAGAT 2280
Db 2112 TATATGACTGGAAATCTTAAAGGTCTTCTTGTGAAAGACAGCCCTCTCTGCTGCCAAGAT 2171
Qy 2281 TTCAATTTCAATGCCAGTTTGTAAAGATTTCTTCCAGATGGAGAAAGGAGTGTCTGTCCA 2340
Db 2172 TTCAATTTCAATGCCAGTTTGTAAAGATTTCTTCCAGATGGAGAAAGGAGTGTCTGTCCA 2231
Qy 2341 TGCAACAGATTTCTCTGTACTTTTAAAGGTGAGCAAGCCCTGTGCTCTGAGGAGAGA 2400
Db 2232 TGCAACAGATTTCTCTGTACTTTTAAAGGTGAGCAAGCCCTGTGCTCTGAGGAGAGA 2291
Qy 2401 TTGGCCAAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGAGAAAGATGCAAG 2460
Db 2292 TTGGCCAAATATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGAGAAAGATGCAAG 2351
Qy 2461 GCATGATTTTATCCAAACCTCGGACGAAACCAAGCTCTCTGCTGATTCGATCAACTGGATC 2520
Db 2352 GCATGATTTTATCCAAACCTCGGACGAAACCAAGCTCTCTGCTGATTCGATCAACTGGATC 2411
Qy 2521 GTGAACAGTTCAACCCCGATGTGAATTAATTTTCCGATATTCGTCTCACTTTGGGATACGCC 2580
Db 2412 GTGAACAGTTCAACCCCGATGTGAATTAATTTTCCGATATTCGTCTCACTTTGGGATACGCC 2471
Qy 2581 CTGCAAGTTGAGTTATGCGAGGAGCCACAGTACCAAAACCTGTGGAAGAGTTATGTGA 2640
Db 2472 CTGCAAGTTGAGTTATGCGAGGAGCCACAGTACCAAAACCTGTGGAAGAGTTATGTGA 2531
|||||

QY 2641 AACTTGGCCACCTCTTAGCAATAGTCCCAAAAGTCAACCAAACTGACAAACAGAGCTGG 2700
Db |||||||
QY 2532 AACTTGGCCACCTCTTAGCAATAGTCCCAAAAGTCAACCAAACTGACAAACAGAGCTGG 2591
Db |||||||
QY 2701 CACAGAGGAGGAAGCCCTCCAAAATAACGGCAGAGAATAAATGAGACAGAGTA 2760
Db |||||||
QY 2592 CACAGAGGAGGAAGCCCTCCAAAATAACGGCAGAGAATAAATGAGACAGAGTA 2651
Db |||||||
QY 2761 CGGTGAGCTAGTAGCAAGATTCTGAAAATCTGGCAATCGCTTCTGATGCTGTCTGAGC 2820
Db |||||||
QY 2652 CGGTGAGCTAGTAGCAAGATTCTGAAAATCTGGCAATCGCTTCTGATGCTGTCTGAGC 2711
Db |||||||
QY 2821 ATGCAATGATGCTAGTCTGTTCTGACCCATCATATCGCTACCAACCAATCCCTTAATGCAAT 2880
Db |||||||
QY 2712 ATGCAATGATGCTAGTCTGTTCTGACCCATCATATCGCTACCAACCAATCCCTTAATGCAAT 2771
Db |||||||
QY 2881 TGGCAAGTTGATAGGATATCTTTCCAAAGATCGTTGTCTGTGACAGCTGGCCATGACTC 2940
Db |||||||
QY 2772 TGGCAAGTTGATAGGATATCTTTCCAAAGATCGTTGTCTGTGACAGCTGGCCATGACTC 2831
Db |||||||
QY 2941 ATCCAAGTCAATTTAAATTTTGGAAATCTTGATCATGCGCAGGAATTCATTTATCTA 3000
Db |||||||
QY 2832 ATCCAAGTCAATTTAAATTTTGGAAATCTTGATCATGCGCAGGAATTCATTTATCTA 2891
Db |||||||
QY 3001 ACTGTGGAATTCGGCAGCCCAATACGGGACAGAGAAAGTTTCATCATGTCACATGCGGA 3060
Db |||||||
QY 2892 ACTGTGGAATTCGGCAGCCCAATACGGGACAGAGAAAGTTTCATCATGTCACATGCGGA 2951
Db |||||||
QY 3061 AGAAGGGATTAAACCTTTGATAATATCATGTGACGCTTGGCCAAAGATGACCAACTC 3120
Db |||||||
QY 2952 AGAAGGGATTAAACCTTTGATAATATCATGTGACGCTTGGCCAAAGATGACCAACTC 3011
Db |||||||
QY 3121 CTTCCAGGATTAAACCAATGACGGTGGAAATCTCGGTGATGCTGTTGTGAATTC 3180
Db |||||||
QY 3012 CTTCCAGGATTAAACCAATGACGGTGGAAATCTCGGTGATGCTGTTGTGAATTC 3071
Db |||||||
QY 3181 TGACCAAGCTGCAATTTGTAATTTGTTTCTTCTAGTCTGGAAGAGAGGATGACCACT 3240
Db |||||||
QY 3072 TGACCAAGCTGCAATTTGTAATTTGTTTCTTCTAGTCTGGAAGAGAGGATGACCACT 3131
Db |||||||
QY 3241 ATCGGACTGCATTTGTTAGCAATGACACCTTGCCATGCTAGCAAGAACTTGAACTGG 3300
Db |||||||
QY 3132 ATCGGACTGCATTTGTTAGCAATGACACCTTGCCATGCTAGCAAGAACTTGAACTGG 3191
Db |||||||
QY 3301 ATCCATTTATGCTGATGCTCAGGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3360
Db |||||||
QY 3192 ATCGATTTATGCTGATGCTCAGGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATG 3251
Db |||||||
QY 3361 CAATGCCAATTTTGAAGCGTTTAATAGAGCTGTTTACTTGGAGGAGGAGCTGGAGG 3420
Db |||||||
QY 3252 CAATGCCAATTTTGAAGCGTTTAATAGAGCTGTTTACTTGGAGGAGGAGCTGGAGG 3311
Db |||||||
QY 3421 AAGCCAGAGATTATTTGGAGCTGCTCTTTAATGATCGGACCTGCGGAAGTCTGGC 3480
Db |||||||
QY 3312 AAGCCAGAGATTATTTGGAGCTGCTCTTTAATGATCGGACCTGCGGAAGTCTGGC 3371
Db |||||||
QY 3481 TCAATTTATCTCTCCACCACTCCAACTACAAGAGCAATACTGATGACCACTTTATTG 3540
Db |||||||
QY 3372 TCAATTTATCTCTCCACCACTCCAACTACAAGAGCAATACTGATGACCACTTTATTG 3431
Db |||||||
QY 3541 AAACTTTCTCAGTTCTACAAAACTTACTAGTTTGAAGCAATGAGGAACTTTTATTTA 3600
Db |||||||
QY 3432 AAACTTTCTCAGTTCTACAAAACTTACTAGTTTGAAGCAATGAGGAACTTTTATTTA 3491
Db |||||||
QY 3601 CTATGTTCTGAGCTTGGCAGGGCATTCATTTGAGAACTGTTGGATTTAAACCATCTGA 3660
Db |||||||
QY 3492 CTATGTTCTGAGCTTGGCAGGGCATTCATTTGAGAACTGTTGGATTTAAACCATCTGA 3551
Db |||||||
QY 3661 CCTTAGGCCAATTCAGAGAAATGGAATTTCTTAGTGACTCCATAATGCAACTGGTAGCCA 3720
Db |||||||
QY 3552 CCTTAGGCCAATTCAGAGAAATGGAATTTCTTAGTGACTCCATAATGCAACTGGTAGCCA 3611
Db |||||||

QY 3721 CAGAGTACTTATTCATTCCATTTCCAGATCATCATGAAAGGACACTTAATCTTTGTTCGAA 3780
Db |||||||
QY 3612 CAGAGTACTTATTCATTCCATTTCCAGATCATCATGAAAGGACACTTAATCTTTGTTCGAA 3671
Db |||||||
QY 3781 GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGGTAGCGAGAGCTGGCATGCAAGGAGT 3840
Db |||||||
QY 3672 GCTCTTTGGTGAATTAATAGAACTCAGGCCAAGGTAGCGAGAGCTGGCATGCAAGGAGT 3731
Db |||||||

RESULT 7
BC088999 4479 bp mRNA linear ROD 27-JUN-2005
BC088999 Mus musculus ribonuclease III, nuclear, mRNA (CDNA clone MGC:115770
IMAGE:6417566), complete cds.
DEFINITION
ACCESSION BC088999
VERSION BC088999.1 GI:57242919
SOURCE MGC.
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 4479)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Sheets,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,J., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4479)
NIH MGC Project
Direct Submission
Submitted (06-JAN-2005) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
Fisher,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
Schetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
Casavant,T., Soares,M.B.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source

Location/Qualifiers
1. .4479
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:115770 IMAGE:6417566"
/tissue_type="Brain, enriched mouse brain 12.5dpc"
/clone_1lb="NIH_BMAP_F10"
/lab_host="DH10B"
/note="Vector: pTX-ASC"
1. .4479
/gene="Rnasen"
/note="synonym: Rn3"
/db_xref="GeneID:14000"
/db_xref="MGI:1261425"
215. .4336
/gene="Rnasen"
/codon_start=1
/product="Rnasen protein"
/protein_id="AAH8999.1"
/db_xref="GI:57242920"
/db_xref="GeneID:14000"
/db_xref="MGI:1261425"
/translation="MGNTCHRMVHPRGCPREGHGARPSPAPFRPNLRLHPQ
OPPAQYQEPSPSPSSYNSQAPSPMPRPDPVPYPPAPSAQGLPFCVPRPYP
NQMRRPFPVPCFPMPMPMPNPPASGAPPGQGTFFWVPPSPMPHPPPVPP
QYNYQYPPGSHSPFPFNSYONSSSPFPSSSSTPHFRHLPYSLPKAQNERR
SPERLKHDDHRRSHGGRHRRSLERRGRSPERRRPSRYSYLDGRGTPPR
HRSYRSDRRHRHRRARRSPSLERSVKKEYKSRYSALPVAPBAGCTPELPG
EMLKTTESWAPPNNVHNSPSREKRAWEKDRWSDSQSGKEKNYTSIKEKAE
EVPKTEBEEELLKPWIRCTHSESYSSDPMQDQSDSTVVGTSRLKLDYDKFEE
LGNRQKAAALPWPFPFKLDDLESSSECEDDTSCSSSDSVDFVIAEIK
RKAHPDRDLDELWPMQMDPLCKCSAKARTGIRHSIYPGEAIKPCRPMTNN
AGLPHYRITVSPNTFLDRTPIVSDHDEYIFEGFSMFAHPLNIPCKVIRFNI
DYTHFIEBMPNPKVGLFLFLFLFLFLFLFLFLFLFLFLFLFLFLFLFLFLFL
RVFRLPDGCKEVLVSHQILLVLRCSKALVPEEIANMLOEELSWQYAECKMI
VTPNGKPSVRIDQDRQFNEVITFFIIVHFGIRPAQLSTAGDQFYQKWLKSTVK
LRHLANSFKVQTKOKLAQBEALQIRKNTMRREVTVLSSQGFQKWTGIRSDVC
QHAMLPVLTIRYHQKLVHMMRKKGINTLINIMSLGODDPTPSRINERLEFLGD
SLNSCGIROPKYGDVKVHMMRKKGINTLINIMSLGODDPTPSRINERLEFLGD
AVVEFTSLVHLYFLPSLEGGLATYRTAIVONHMLAKLELDRLFMVLAHGPDL
RESDLRHMANECALIGAVLESLBEAKQLFGRLLFNDPDLREVWLNYPHLPLQ
EPNTDQLLETSPVLQKLEFEAEIGVIFTHVRLRLARFTLRVTGPNHLLTGLHNRME
FLGDSIMQVLTATLLESTIAALYIDKLEYVHTFMNVCFPRLEKFIINQWNPDKS
TKRPVALRKTLLADLESIAALYIDKLEYVHTFMNVCFPRLEKFIINQWNPDKS
OLQCCCLTLRTEGKEPDIPLYKTLOTVPSPHARTYVAVFKGERIGCGKGPSIQAB
MGAAMDALKYNFPQAHQKRFYERKYRQELKEMRWERHQREREPEAEIDKK"

ORIGIN

Query Match 70.9%; Score 3379.6; DB 9; Length 4479;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 3762; Conservative 0; Mismatches 574; Indels 8; Gaps 3;
QY 151 GGAAGAGTTAGAGTGTGATCTTCACTATGATATGAGGAGTCTCTGAGCTTATATTC 210
DB 117 GGAGGAGTTCAACATGGGACACACGACTGTCAAGCGGAGTCTTGAACTCATCG 176
QY 211 TCTGTGGAAGATGTGACATATCAGCGGAAACATCATGATGACAGGGAACACATGTGACA 270
DB 177 TGAGCGGAAGATGTAGCTTTTCTACAGGACGATCATGATGCAAGCAATATCTGTGATA 236
QY 271 GAAATGTCGTTTCAACCGGGACGAGGGCGTCCCGGAGGACGAGGAGGACATGAGCCAGAC 330
DB 237 GAATGTCGTACCAACCCAGGACGAGGGTGTCCCGGGGCGGAGGAGGACGAGGCGCAGAC 296
QY 331 CCTCAGACCATCTCTTTAGGCCCCAAAATCTGAGGCTGTCTTACCTCAGCAGCCTCTCTG 390
DB 297 CTTTCAAGCACCAGCTTTTCAAGACCCCAAAACCTTGCAGCTTCTTATCCCTCCAGCAGCCGCTG 356
QY 391 TGCATATATCAATATGAATCACTCCAGTGGCCCTTCCACCACTTCTCAAACTCTCCAGGCC 450
DB 357 CGCAGTATCAATATGAGCTTCCAGCGCCGCTCTTCTCTCTACTCGAACTCTCAGGCC 416

[illegible]

Db	3654	TTCTGGCGCAGCTCCATAATGCAGCTGGTGGCCACAGAGTACTTGTTCATTCATTTCCCGG	3713
QY	3748	ATCATCATGAAGAGCACCTTAACTTTGTGTGGAGAGCTCTTTGGTGAATAATAGAACTCAGG	3807
Db	3714	ACCATCAGCAAGGACACATTCGACGTGTGTACGAAGTTCCTTAGTGAACAAACAGAACGCAAG	3773
QY	3808	CCAAAGGTAGCGGAGGAGCTGGGCATGCAGAGAGTACGCCATAACCAACGACGAACCAAGA	3867
Db	3774	CCAAAGGTAGCAGAGAACTGGGCATGCAGAGATATGCCATCATCAACGACGAACCAAGA	3833
QY	3868	GGCCTGTGGCGCTTCGCCACCAAGACCTTGGCGGACCTTTTGGGAATCATTTATTGCAAGCGC	3927
Db	3834	GACCCGTGGCCCTAAAGAACCAGACTTTGGCAGACCTTTTGGGAATCATTTATGCGAGCGC	3893
QY	3928	TGTACACTGATTAAGATTTGGGAATATGTTTCATATCTTTTCATGAAATGTCTGCTTTTCCAC	3987
Db	3894	TGTACACTCGAACAAGGACCTGGGAATATGTCCACACTTTTCATGAACGCTCTGCTTTTCCCC	3953
QY	3988	GATTGAAGAAATTCATTTTGAATCAGGATTTGGAATGACCCCAATCCAGCTTCAGCAGAT	4047
Db	3954	GGCTGAAGGAGTTCATTTCTGAATCAGGATTTGGAACGCCCAAGTCGCAGCTTCGACGAGT	4013
QY	4048	GTTGCTTTGACCTTAGCAGACGAAGGAAAAGAGCGCAGACATTCCTCTGTACAAGACTCTGCG	4107
Db	4014	GTTGCTCTGACCTTGAGGACGAAGGGAAGAGCGCTGACATCCCTTATACAAGACTCTGCG	4073
QY	4108	AGACAGTGGGCCCCATCCCATGCCCGAACCTACACTGTGGCTGTTTATTTCAAGGGAGAAA	4167
Db	4074	AGACAGTGGGCCCCATCCCATGCTAGAACCTTACACTGTGGCTGTTTATTTCAAGGGAGAAA	4133
QY	4168	GAATAGGCTGTGGGAAGGACCAAGTATTCAGCAAGCGGGAATTTGGGAGCAGCAATGGATG	4227
Db	4134	GGATAGGCTGTGGGAAGGACCAAGCAATTCAGCAGGCGGAGATGGGAGCAGCAATGGATG	4193
QY	4228	CGCTTCAAAAATAATATTTTCCCACAGATGCCCATCAGAAAGCGGTTTCATCGAACGGAAGT	4287
Db	4194	CAGTGAAGAAATATACTTTTCCCAGATGGCCCATCAGAAAGCGGTTTCATTTGACGCGGAAT	4253
QY	4288	ACAGACAGAGTTTAAAGAAATAGGTGGGAAAGAGAGAGCATCAAGAGAGAGAGCCAGATG	4347
Db	4254	ACAGACAGAGTTTAAAGAAATAGGTGGGAAAGAGAGAGCATCAGGAGAGAGAGCCGAGG	4313
QY	4348	AGACTGAAGACATCAAGAAATAAGAGAGGCGATGCAAGTGTGGAGTATTTACTTGTCTCAG	4407
Db	4314	AGGCTGAAGACATCAAGAAATAACGAGGCGAGG--AGTGGCGGAGCAITTGCCCTTTCTAA	4371
QY	4408	TAACTGTGACTGTGTCTATTGAGACCTAGCCTAGTATTTCCTGCAGCAATGAACGAGT	4467
Db	4372	GAACTGTGACTGTGGCCCATCGAGACTAGCCTGGTTTTTCCCTTAGACAATGAATGAAGT	4431
QY	4468	GTGCTCATTTGAATAAATAACAGA	4491
Db	4432	GTGCCCATTTGAATAAATAACTCAA	4455
RESULT 8			
LOCUS	BC060265		
DEFINITION	Mus musculus ribonuclease III, nuclear, mRNA (cdna clone IMAGE:5698108), partial cds.	linear	ROD 27-JUN-2005
ACCESSION	BC060265		
VERSION	BC060265.1		
KEYWORDS	GI:37994703		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 3838)		
AUTHORS	Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,		

DIATCHENKO, L., MARUSINA, K., FARMER, A.A., RUBIN, G.M., HONG, L., SCAPLETEN, M., SOARES, M.B., BONALDO, M.F., CASAVANT, T.L., SCHLETZ, T.E., BROWNSTEIN, M.J., USHDIN, T.B., TOSHIYUKI, S., CARRINCI, P., PRANGE, C., RAHA, S.S., LOQUELLANO, N.A., PETERS, G.J., ABRAMSON, R.D., MULLAHY, S.J., BOSAK, S.A., MCEWAN, P.J., MCKERNAN, K.J., MALEK, J.A., GUNARATNE, P.H., RICHARDS, S., WORLEY, K.C., HALE, S., GARCIA, A.M., GAY, L.J., HULYK, S.W., VILLALON, D.K., MUZNY, D.M., SODERGREN, E.J., LU, X., GIBBS, R.A., FAHEY, J., HELTON, E., KETTEMAN, M., MADAN, A., RODRIGUES, S., SANCHEZ, A., WHITING, M., MADAN, A., YOUNG, A.C., SHEVCHENKO, Y., BOUFFARD, G.G., BLAKESLEY, R.W., TOUCHMAN, J.W., GREEN, E.D., DICKSON, M.C., RODRIGUEZ, A.C., GRIMWOOD, J., SCHMUTZ, J., MYERS, R.M., BUTTERFIELD, Y.S., KRZYWINSKI, M.I., SKALSKA, U., SMALUS, D.B., SCHMERCH, A., SCHEIN, J.E., JONES, S.J. AND MARRA, M.A.			
CONSRMTM TITLE	Mammalian Gene Collection Program Team Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
REFERENCE	12477932		
AUTHORS	2 (bases 1 to 3838)		
CONSRMTM TITLE	NIH MGC Project Direct Submission		
JOURNAL	Submitted (20-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA		
REMARK COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabbs@mail.nih.gov Tissue procurement: Dr. Jim Lin, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Martin Hirst, Thomas Zeng, Ryan Morin, Michelle Moksa, Johnson Pang, Diana Mah, Jing Wang, Kieth Fichter, Eric Chuah, Allen Delaney, Rob Kirkpatrick, Agnes Baross, Sarah Barber, Mabel Brown-John, Steve S. Chand, William Chow, Ryan Babakaiff, Dave Wong, Corey Matsuo, Jaclyn Beland, Susan Gibson, Luis delRio, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Mike R. Mayo, Josh Moran, Diana Palmquist, JR Santos, Duane Smalus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Steven Jones, Rob Holt, Marco Marra.		
FEATURES	Location/Qualifiers		
source	1..3838 /organism="Mus musculus" /mol_type="mRNA" /strain="CS7BL/6" /db_xref="taxon:10090" /clone="IMAGE:5698108" /tissue_type="Brain, enriched mouse brain 12.5dpc" /clone_lib="NIH EMAP_F10" /lab_host="DH10B" /note="Vector: pYX-ASC" c1..3838 /gene="Rnaase" /note="synonym: Rn3" /db_xref="GeneID:14000" /db_xref="MGI:1261425" c1..3697 /gene="Rnaase" /codon_start=2 /product="Rnaase protein" /protein_id="AAH60265.1" /db_xref="GI:37994704" /db_xref="GeneID:14000"		
gene			
CDS			

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK plate: 132 Row: a Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

KEYWORDS	JP 2002191363-A/10759.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 2598)
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE	Primer for synthesizing full-length cDNA and use thereof
JOURNAL	Patent: JP 202191363-A 10759 09-JUL-2002;
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/10759 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO, PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10, C12P21/02, C12Q1/68/ C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT CDS (10) . . (2316) . Location/Qualifiers source 1. .2598 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
ORIGIN	
Query Match	54.2%; Score 2582; DB 6; Length 2598;
Best Local Similarity	99.6%; Pred. No. 0;
Matches 2588; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
Qy	2052 GGATTTCTATGTTTGACATGCCCCCTGACCAATATTCACATGCTGTAAAGTAATTAGA 2111
Db	
Qy	2112 TTCAACATAGACTACAGATTCAATTCATTGAAGAGATGATGCCGGAGAAATTTTGTGTG 2171
Db	
Qy	2172 AAAGGGCTGAACCTCTTTTCACTGTCTATTCAGAGATATTTTGGAAATATATGACTGG 2231
Db	
Qy	2232 AATCTTAAAGTCTTTGTTTGAACAGACGCCCTCCCTGCTGCCCAAGATTCATTTCATG 2291
Db	
Qy	2292 CCACGTTTTGTAAGATTCTTCCAGATGGAGAAAGAGTGCTGTCCATGCACCAAGATT 2351
Db	
Qy	2352 CTCCTGTACTTTAAAGTGCAGCAAGCCCTGTGTGCTGAGGAGGAGATTGCCAAATG 2411
Db	
Qy	2412 CTTAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAGAGAAATGCNAGGCATGATTGT 2471
Db	
Qy	2472 ACCAACCTCGGACGAAACCAAGCTCTGTCGTATCGATCAACTGGATCGTGAACAGTT 2531
Db	
Qy	2532 AACCCCGATGTATTCATTTTCGATATCGTCCACTTTGGGATACGCCCTGCACAGTTG 2591
Db	
Qy	2592 AGTTATGCGAGGAGACCCACAGTACCAAAAACTGTGGAAGAGTTATGTATGTGAAACTTTGCGCAC 2655
Db	
Qy	2652 CTCTAGCAAAATAGTCCCAAGTCAACAAACTGACAAAACAGAAAGCTGGCCACAGAGGGAG 2711
Db	
Qy	2712 GAAGCCCTCCAAAAATACGCGCAGAGAATACAAATGAGACGAGAAAGTAACGCTGGAGCTA 2771
Db	
Qy	2772 AGTAGCCAAAGGATTTCTGGAAAACTGGCATCCGTTCTGTCTCTGATGTCGTGATGCAATGATG 2831
Db	
Qy	2832 CTACCTGTTCTGACCCATCATATCCGCTACCAACAATGCCATAATGCATTTGGAGCAAGTTG 2891
Db	
Qy	2892 ATAGGATATACTTTTCAAAGATCGTGTCTGTGAGCTGGCCCATGACTCATCAAAGTCAT 2951
Db	
Qy	2952 CATTTAAATTTTGGAAATGAATCCTGATCATGCCAGGAATTCATTATCTAACTGTGGAATT 3011
Db	
Qy	3012 CGGAGGCCAAAATACGGAGACAGAAAAGTTTCATCATGACATGCGGAGAGAAAGGAT 3071
Db	
Qy	3072 AACACCTTGATAAATAATCATGTCAGCTGGCCCAAGATCACCAACTCCCTCGAGGATT 3131
Db	
Qy	3132 AACCACAAATGAACGGTTTGGAAATTCCTGGGTGATGCTGTGTTGAAATTTCTGACCAAGCTC 3191
Db	
Qy	3192 CATTTGATCTATTTGTTTCTAGTCTGGAAGAGGAGGATTAGCAACCTATCGGACTGCC 3251
Db	
Qy	3252 ATTTGTTAGAAATCAGACACCTTGGCCATCTAGCAAAAGAACTTGAACCTGGAATCGAATTTATG 3311
Db	
Qy	3312 CTGTATGCTCAACGGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATGCAATGGCCAAAT 3371
Db	
Qy	3372 TGTTTTGAACGGTTAATAGAGCTGTTTACTTGGAGGAAAGCCCTGGAGGAAGCCCAAGCAG 3431
Db	
Qy	3432 TTATTTGGAGCGCTTGTCTTTTAATGATCCGGACTTCGCGAAGTCTGGCTCAATTAATTCCT 3491
Db	
Qy	3492 CTCACCCACTCCAACTTACAAGAGCCAAATACATGATCGACAACTTAATTGAAACTTCTTCCCA 3551
Db	
Qy	3552 GTTCTACAAAACTTACTGAGTTTGAAGAGCAATTGGAGTAATTTTATCTCATGTTCGA 3611
Db	
Qy	3612 CTTCTGCAAGGGCAATTCATTTGAGNACTGTGGGATTTAAACCATCTGACCCCTAGGCCAC 3671
Db	
Qy	3671 CTTCTGCAAGGGCAATTCATTTGAGNACTGTGGGATTTAAACCATCTGACCCCTAGGCCAC 3731
Db	

|||||
421 ACCAACCTGGGAGCAACAGCTCTGTCGGTATCGATCAACTGGATGTAACAGTTTC 480
QY
2532 AACCCGATGTGATTAATCTTTCCGATATCGTCCACTTTGGGATACGCCCTGGACAGTTTG 2591
Db
481 AACCCGATGTGATTAATCTTTCCGATATCGTCCACTTTGGGATACGCCCTGGACAGTTTG 540
QY
2592 AGTTATGCAAGGAGACCCACAGTACCAAAACCTGTGGAAGAGTTATGTGAAACTTCGCCAC 2651
Db
541 AGTTATGCAAGGAGACCCACAGTACCAAAACCTGTGGAAGAGTTATGTGAAACTTCGCCAC 600
QY
2652 CTCCTAGCAAAATAGTCCCAAGTCAAAACCTGACAAAACAGAGCTGGCAACAGAGGGAG 2711
Db
601 CTCCTAGCAAAATAGTCCCAAGTCAAAACCTGACAAAACAGAGCTGGCAACAGAGGGAG 660
QY
2712 GAAGCCCTCCAAAAATA CGGCAAGAAATA CAATGAGA CGAGAAATACGGTGGAGCTA 2771
Db
661 GAAGCCCTCCAAAAATA CGGCAAGAAATA CAATGAGA CGAGAAATACGGTGGAGCTA 720
QY
2772 AGTAGCCRAGGATCTGGAAACTGGCATCCGTTCTGATGTCTGTCAGCATGCAATGATG 2831
Db
721 AGTAGCCRAGGATCTGGAAACTGGCATCCGTTCTGATGTCTGTCAGCATGCAATGATG 780
QY
2832 CTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCCTAATGCATTTGGACRAGTTG 2891
Db
781 CTACCTGTTCTGACCCCATCATATCCGCTACCAACCAATGCCTAATGCATTTGGACRAGTTG 840
QY
2892 ATAGGATATACCTTTCCAGATCGTTCTGTGTGACGTGGCCATGATCATCAAGTCAT 2951
Db
841 ATAGGATATACCTTTCCAGATCGTTCTGTGTGACGTGGCCATGATCATCAAGTCAT 900
QY
2952 CATTTAAATTTGGAAATGAATCCTGATCATGCCAGGAATTCATTATCTAACTGTGGAAT 3011
Db
901 CATTTAAATTTGGAAATGAATCCTGATCATGCCAGGAATTCATTATCTAACTGTGGAAT 960
QY
3012 CGSCACCCAAATACGGAGACAGAAAGTTTCATCATGCAATGCGGAGAAAGGGAAT 3071
Db
961 CGSCACCCAAATACGGAGACAGAAAGTTTCATCATGCAATGCGGAGAAAGGGAAT 1020
QY
3072 AACACCTTGATTAATATCATGTCAACGCTTTGGCCAAAGATGAACCACTCCCTCGAGGAT 3131
Db
1021 AACACCTTGATTAATATCATGTCAACGCTTTGGCCAAAGATGAACCACTCCCTCGAGGAT 1080
QY
3132 AACCAATGAAGCGTTGGAAATTCCTGGTGATGCTGTGTTGAAATTTCTGACCGCTC 3191
Db
1081 AACCAATGAAGCGTTGGAAATTCCTGGTGATGCTGTGTTGAAATTTCTGACCGCTC 1140
QY
3192 CATTTGCTACTATTTGTTCTAGTCTGGAAGAGGAGGATTAGCAACCTATCGACTGCC 3251
Db
1141 CATTTGCTACTATTTGTTCTAGTCTGGAAGAGGAGGATTAGCAACCTATCGACTGCC 1200
QY
3252 ATTTGTTCAGAAATCAGCACTTGCCATGCTAGTCAAAAGAAACTTGAACTGGATCCATTTATG 3311
Db
1201 ATTTGTTCAGAAATCAGCACTTGCCATGCTAGTCAAAAGAAACTTGAACTGGATCCATTTATG 1260
QY
3312 CTGTATGCTCAGGGGCTGACCTTTGTAGGAATCGGACTTCGACATGCAATGGCCAAAT 3371
Db
1261 CTGTATGCTCAGGGGCTGACCTTTGTAGGAATCGGACTTCGACATGCAATGGCCAAAT 1320
QY
3372 TGTGTTGAGCGTTAATAGGAGCTGTTTACTTGGAGGAGCCCTGGAGGAAGCCCAAGCAG 3431
Db
1321 TGTGTTGAGCGTTAATAGGAGCTGTTTACTTGGAGGAGCCCTGGAGGAAGCCCAAGCAG 1380
QY
3432 TTAATTTGAGCGTTGCTCTTTAATGATCGGACCTCGCGAAGTCTGGCTCAATTTACCT 3491
Db
1381 TTAATTTGAGCGTTGCTCTTTAATGATCGGACCTCGCGAAGTCTGGCTCAATTTACCT 1440
QY
3492 CTCACCCACTCAACTA CAAAGGCCAAATAC TGTATCGCA CAACCTTAT TGAACCTTCCTCA 3551
Db
1441 CTCACCCACTCAACTA CAAAGGCCAAATAC TGTATCGCA CAACCTTAT TGAACCTTCCTCG 1500
QY
3552 GTTCTCAAAAATCTTACTGAGTTTGAAGAGCAATTTGGAGTAATTTTACTCATGTTTCA 3611
|||||

1501 GTTCTACAAAACTTTACTGAGTTTGAAGAAGCAATTTGGAGTAATTTTACTCATGTTCA 1560
QY
3612 CTTCTGGCAAGGGCAATTCACATTTGAGAACTGTGGCAATTTAAACCATCTGACCCCTAGGCCAC 3671
Db
1561 CTTCTGGCAAGGGCAATTCACATTTGAGAACTGTGGCAATTTAAACCATCTGACCCCTAGGCCAC 1620
QY
3672 AATCAGAGAAATGGAATTCCTAGTGACTCCATAATTCGCAACTGCTGTAGCCACAGACTCTTA 3731
Db
1621 AATCAGAGAAATGGAATTCCTAGTGACTCCATAATTCGCAACTGCTGTAGCCACAGACTCTTA 1680
QY
3732 TTCATTTCAATTTCCCAAGATCATGAGAGCACTTAACCTTTGTTGGGAAGCTCTTTGGTG 3791
Db
1681 TTCATTTCAATTTCCCAAGATCATGAGAGCACTTTAACTTTTGTGCGAAGCTCTTTGGTG 1740
QY
3792 AATAATAGAACTCAGGCCAAAGGTAGCGGAGGAGCTGGGCATCGAGAGTACGCCATAAAC 3851
Db
1741 AATAATAGAACTCAGGCCAAAGGTAGCGGAGGAGCTGGGCATCGAGAGTACGCCATAAAC 1800
QY
3852 AACGCAAGACCAAGAGGCTGTGGCGCTTCGCACCAAGACCTTTGGCGGACCTTTTGGAA 3911
Db
1801 AACGCAAGACCAAGAGGCTGTGGCGCTTCGCACCAAGACCTTTGGCGGACCTTTTGGAA 1860
QY
3912 TCATTTATTTGAGCGCTGTACACTGATAGGAATTTGGAATATGTTCAATCTTTTCATGAT 3971
Db
1861 TCATTTATTTGAGCGCTGTACACTGATAGGAATTTGGAATATGTTTCATCTTTTCATGAT 1920
QY
3972 GTCTGCTTTTCCACGATTTGAAAGAAATTCATTTTGAATCAGGATTTGGAATGACCCCAA 4031
Db
1921 GTCTGCTTTTCCACGATTTGAAAGAAATTCATTTTGAATCAGGATTTGGAATGACCCCAA 1980
QY
4032 TCCAGCTTTCAGCAGCTGTGTTGACACTTAGGACAGAGGAAAGAGCCAGACATTCCT 4091
Db
1981 TCCAGCTTTCAGCAGCTGTGTTGACACTTAGGACAGAGGAAAGAGCCAGACATTCCT 2040
QY
4092 CTGTCAAGACTCTGCGAGACAGTGGGCCCATCCCATGCCCCGAACTTACACTGTGGCTGT 4151
Db
2041 CTGTCAAGACTCTGCGAGACAGTGGGCCCATCCCATGCCCCGAACTTACACTGTGGCTGT 2100
QY
4152 TATTTCAAGGAGAAAGATAGCTGTGGAAAGGACCAAGTATTCAGCAAGCGGAAATG 4211
Db
2101 TATTTCAAGGAGAAAGATAGCTGTGGAAAGGACCAAGTATTCAGCAAGCGGAAATG 2160
QY
4212 GGAGCAGCAATGGATCGCTTGAAAAATAATAATTTTCCCAGATGGCCCATCAGAAAGCGG 4271
Db
2161 GGAGCAGCAATGGATCGCTTGAAAAATAATAATTTTCCCAGATGGCCCATCAGAAAGCGG 2220
QY
4272 TTCATCGAAACGGAAGTACAGACAAGAGTTAAAGAAATGAGGTGGGAAAGAGAGCATCAA 4331
Db
2221 TTCATCGAAACGGAAGTACAGACAAGAGTTGAAAGAAATGAGGTGGGAAAGAGAGCATCAA 2280
QY
4332 GAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGAGGGGCAATGCAAGTGTGGA 4391
Db
2281 GAGAGAGCCAGATGAGACTGAGACATCAAGAAATAAAGAGGGGCAATGCAAGTGTGGA 2340
QY
4392 GTATTTTACTTGTCTCAGTAACTGTGACTGTGTCTATTTGAGACCTAGCTTTTCTCTGC 4451
Db
2341 GTATTTTACTTGTCTCAGTAACTGTGACTGTGTCTATTTGAGACCTAGCTTTTCTCTGC 2400
QY
4452 AGACAATGAACGAGTGTGCTCATTTGAAATAAATAACAGAGTCAAAATCGCTATTGTTGTT 4511
Db
2401 AGACAATGAATGAAGTGTGCTCATTTGAAATAAATAACAGAGTCAAAATCGCTATTGTTGTT 2460
QY
4512 TTAATGATCTGTTTTTGTAGCTGGATGCTTTTATTAATAAGTATTAAGATTTTCTCTTCTATT 4571
Db
2461 TTAATGATCTGTTTTTGTAGCTGGATGCTTTTATTAATAAGTATTAAGATTTTCTCTTCTATT 2520
QY
4572 TTAACGAAAACTGACTTTTGGTGAATGTCATTTACTTCTTTTATTTTGTCTTTAAATA 4631
Db
2521 TTAACGAAAACTGACTTTTGGTGAATGTCATTTACTTCTTTTATTTTGTCTTTAAATA 2580
QY
4632 ATAAAAATTCAGAGCAT 4649
Db
2581 ATAAAAATTCAGAGCAT 2598

RESULT 11	AK001121	2598 bp	mRNA	linear	PRI 30-JAN-2004
LOCUS	AK001121				
DEFINITION	Homo sapiens cDNA FLJ10259 fis, clone HEMBB1000947, highly similar to Homo sapiens clone HAW100 putative ribonuclease III mRNA.				
ACCESSION	AK001121				
VERSION	AK001121.1	GI:7022182			
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.				
	1				
AUTHORS	Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirakawa, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Tanase, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotaka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, M., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hiki, J., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Katsumura, K., Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.				
TITLE	Complete sequencing and characterization of 21,243 full-length human cDNAs				
	Nat. Genet. 36 (1), 40-45 (2004)				
JOURNAL	14702039				
PUBMED	2				
AUTHORS	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.				
	NEDO human cDNA sequencing project				
TITLE	Unpublished				
	3 (bases 1 to 2598)				
JOURNAL	Isogai, T. and Otsuki, T.				
REFERENCE	Direct Submission				
AUTHORS	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1539-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)				
	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3' - end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.				
FEATURES	Location/Qualifiers				

source	1. .2598				
	/organism="Homo sapiens"				
	/molecule="mRNA"				
	/db_xref="taxon:9606"				
	/clone="HEMBB1000947"				
	/issue type="whole embryo, mainly body"				
	/clone lib="HEMBB1"				
	/dev stage="embryo, 10 weeks"				
	/note="cloning vector: pME18SFL3"				
	10. .2319				
	/note="unnamed protein product"				
CDS	/codon start=1				
	/protein id="BAA91511.1"				
	/db_xref="GI:7022183"				
	/translation="MPAHPALNIPLCVKVIRFNIDYTHPIBEMMPENPCVKGLELFS LFLFDILLEYDNLNKLQFEDSPCCPRFEMPRFVFLPDGKGVLSHQHQLKLL RSCSKALPEEIANLQWEELEWQYAECKGMIVTNPGTKSSVDRQDQRFQNP ALQTKRQNDREVTVELSSQGWKTGIRSDVCOHAMLLPVLTHIRVHQCLMLDK LGIYTFODPCLLQAMTHPSHLNFGMNPDHARNSLNCGIROPKYDRKRVHMRK KGINTLINMSRLGQDDPTPSRINENRLEPLGDVVEFLTSVHLXYLPPSLEGLA TYRTAIVQNHMLAKLELDRFMDYAHGPDLCRESLDURHAMWANGFEALIGAVYLEG SLSEAKQLFGRLLFNDPDLREVLYNPLHPLQLEPNTDRLIETSPVLOKTFESEA IGVIETHVRLARAFTRTVGFNHLTLGHNQRMFLGDSIMQLVATYFLFIHFDHEE IDKDLRYVHTFMNVCFEPLEKFEILNQDNDPKSQOCCLTLTETGKSPDIPLYKTL QTVGSHARTYTVAVYFKGERIGCGKGPSIQQAEMGAMDALEKYNFPQMAHQKRFIE RKYRQELKEMRWERHREPEDETDIKK"				
ORIGIN					
	Query Match	54.2%;	Score 2582;	DB 8;	Length 2598;
	Best Local Similarity	99.6%;	Pred. No. 0;		
	Matches 2588;	Conservative	0;	Mismatches	10;
				Indels	0;
				Gaps	0;
QY	2052	GGATTTTTCATGTTGTCACATGCCCCCTGACCAATATTCACATGTTGTAAGTAATAGG	2111		
DB	1	GGATTTTTCATGTTGTCACATGCCCCCTGACCAATATTCACATGTTGTAAGTAATAGG	60		
QY	2112	TTCAACATAGACTACACGATTCATTTTCATGATGATGATGATGATGATGATGATGATG	2171		
DB	61	TTCAACATAGACTACACGATTCATTTTCATGATGATGATGATGATGATGATGATG	120		
QY	2172	AAAGGCTTGAACCTCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATG	2231		
DB	121	AAAGGCTTGAACCTCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATG	180		
QY	2232	AATCTTAAAGGTCCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATG	2291		
DB	181	AATCTTAAAGGTCCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATG	240		
QY	2292	CCACGTTTGTGAAGATTTCTTCAGATGAGGAGGAGGAGGATGATGATGATGATGATG	2351		
DB	241	CCACGTTTGTGAAGATTTCTTCAGATGAGGAGGAGGAGGATGATGATGATGATGATG	300		
QY	2352	CTCCTGTACTTCTTAAGTGCGACGAGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGG	2411		
DB	301	CTCCTGTACTTCTTAAGTGCGACGAGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGG	360		
QY	2412	CTTTCAGTGGGAGGAGCTGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2471		
DB	361	CTTTCAGTGGGAGGAGCTGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	420		
QY	2472	ACCAACCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2531		
DB	421	ACCAACCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	480		
QY	2532	AACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2591		
DB	481	AACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	540		
QY	2592	AGTTATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2651		
DB	541	AGTTATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	600		

Qy	2652	CTCCTAGCAAAATAGTCCCAAAGTCAAAACAACTGACAAACAGAAAGCTGGCAACAGAGGAG	2711
Db	601	CTCCTAGCAAAATAGTCCCAAAGTCAAAACAACTGACAAACAGAAAGCTGGCAACAGAGGAG	660
Qy	2712	GAAGCCCTCAAAAATACGCGAAGAATACAAATGAGACGAGAGTAACGGTGGAGCTA	2771
Db	661	GAAGCCCTCAAAAATACGCGAAGAATACAAATGAGACGAGAGTAACGGTGGAGCTA	720
Qy	2772	AGTAGCAAGGATTCTGGAATACTGCGATCCGTTCTGATGTCGTGACAGATGCAATGATG	2831
Db	721	AGTAGCAAGGATTCTGGAATACTGCGATCCGTTCTGATGTCGTGACAGATGCAATGATG	780
Qy	2832	CTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCCTAAATGCAATTTGGACAAGTTG	2891
Db	781	CTACCTGTTCTGACCCATCATATCCGCTACCAACCAATGCCTAAATGCAATTTGGACAAGTTG	840
Qy	2892	ATAGGATATACCTTTCAAAGATCGTTGTCGTTGAGCTGGCCATGATCATATCCAAAGTCAT	2951
Db	841	ATAGGATATACCTTTCAAAGATCGTTGTCGTTGAGCTGGCCATGATCATATCCAAAGTCAT	900
Qy	2952	CATTTAAATTTTGGAAATGAAATCTGATCATATGCGAGGAATTCATTTCTAACTGGAATTT	3011
Db	901	CATTTAAATTTTGGAAATGAAATCTGATCATATGCGAGGAATTCATTTCTAACTGGAATTT	960
Qy	3012	CGGACGCCAAATACGGAGACAGAAAAGTTTCATCATGTCATGCGGAGAAAGAGGANTT	3071
Db	961	CGGACGCCAAATACGGAGACAGAAAAGTTTCATCATGTCATGCGGAGAAAGAGGANTT	1020
Qy	3072	AACACCTTGATAAATATCATGTGACGCCCTGGCCAAAGATGACCCAACTCCCTCGAGGANTT	3131
Db	1021	AACACCTTGATAAATATCATGTGACGCCCTGGCCAAAGATGACCCAACTCCCTCGAGGANTT	1080
Qy	3132	AACCAAAATGAAAGGTTGGAATTCCTGGTGTATGCTGTTGTTGAATTTCTGACACGCTC	3191
Db	1081	AACCAAAATGAAAGGTTGGAATTCCTGGTGTATGCTGTTGTTGAATTTCTGACACGCTC	1140
Qy	3192	CATTTGCTACTATTGTTTCTTAGTCTGGAAGAAGAGGATTTAGCAACTCTATCGCACTGCC	3251
Db	1141	CATTTGCTACTATTGTTTCTTAGTCTGGAAGAAGAGGATTTAGCAACTCTATCGCACTGCC	1200
Qy	3252	ATTGTTCAGAATCAGACACTTGCCATGCTAGCAAAAGAACTTGAACTGGATTCATTTATG	3311
Db	1201	ATTGTTCAGAATCAGACACTTGCCATGCTAGCAAAAGAACTTGAACTGGATTCATTTATG	1260
Qy	3312	CTGTATGCTCAGGGGCTGACCTTTGTAGAGAAATCGGACCTTTCGACATGCAATGGCCAAAT	3371
Db	1261	CTGTATGCTCAGGGGCTGACCTTTGTAGAGAAATCGGACCTTTCGACATGCAATGGCCAAAT	1320
Qy	3372	TGTTTTGAAAGGTTAATAGAGCTGTTTACTTGGAGGAGAGCCTGGAGGAGCCAGCAG	3431
Db	1321	TGTTTTGAAAGGTTAATAGAGCTGTTTACTTGGAGGAGAGCCTGGAGGAGCCAGCAG	1380
Qy	3432	TTATTTGGACGCTTGCTCTTTTAATGATCGGACCTCGCGGAAGTCTGGCTCAATATATCCT	3491
Db	1381	TTATTTGGACGCTTGCTCTTTTAATGATCGGACCTCGCGGAAGTCTGGCTCAATATATCCT	1440
Qy	3492	CTCCACCCATCCAACTACAAGGCAAAATATGATGACAACTATTTATGAAAATTTCTCCA	3551
Db	1441	CTCCACCCATCCAACTACAAGGCAAAATATGATGACAACTATTTATGAAAATTTCTCCG	1500
Qy	3552	GTTTCTCAAAAATCTTACTAGTTTGNAGAGCAATTTGGAGTAATTTTTTACTCATGTTTCCA	3611
Db	1501	GTTTCTCAAAAATCTTACTAGTTTGNAGAGCAATTTGGAGTAATTTTTTACTCATGTTTCCA	1560
Qy	3612	CTTCTGGCAAGGCAATTCATATGAGAACTGTGGGATTTAAACCATCTGACCCCTAGGCCAC	3671
Db	1561	CTTCTGGCAAGGCAATTCATATGAGAACTGTGGGATTTAAACCATCTGACCCCTAGGCCAC	1620
Qy	3672	AATCAGAGAAATGGAATTCCTAGGTGATCTCCATAATATGCAACTGGTATGACCAAGAGTACTTA	3731
Db	1621	AATCAGAGAAATGGAATTCCTAGGTGATCTCCATAATATGCAACTGGTATGACCAAGAGTACTTA	1680

ORGANISM	Gallus gallus	273	ATGTCGTTCCACCCCGGACGAGGGCGTCCCGAGGACGAGGAGGACATCGAGCCAGACC	332
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.	173	ATGTCGTTTCCATTCTGGACGAGGATGTCCTCCGTCGACGGGGACGAGGAGCAAGAACT	232
AUTHORS	1 Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezzubov, Y., Zaim, J., Fiedler, P., Kutter, S., Blagoderaki, A., Kostovska, D., Koter, M., Plachy, J., Carninci, P., Hayashizaki, Y., and Buerstedde, J.M.	333	TCAGCACCATCTCTTAGGCCCCCAAAATCTGAGGGCTGCTTACACCTCAGCAGCCTCTCTGTG	392
TITLE	Full-length cDNAs from chicken bursal lymphocytes to facilitate gene function analysis	233	TCAGCACAGACGTATCGTCTCTCAGAACCTGCGGCGAGCTTTCACCCCCAGCAGCCTTCTGTA	292
JOURNAL	Genome Biol. 6 (1), R6 (2005)	393	CAATATCAATATGAACCTCCAAGTGGCCCTTTCACCACTTCTTCAAACCTCTCCAGCCCCC	452
PUBMED	15642098	293	CAGTACCAAGTATGACCAAGCCGACCTGACGTACCTATTTCAAAACCTCTTCAGCCACC	352
REFERENCE	2 (bases 1 to 4516)	453	AAATTTCTCTCCACGACGACGACTTTGTAGCTTCCCTCCCTCCCTCCCTCCCTCCGTCAGCG	512
AUTHORS	Caldwell, R.B.	353	AGTTACATACGCTCCAGGCCAGACTTTGTTTCATATCCCCACAGTTCCTCCCTTCCACA	412
TITLE	Direct Submission	513	CAAGGCCCTCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC	572
JOURNAL	Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum, Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr. 1, D-85764 Neuherberg, GERMANY	413	CAAAATCCCATCAACCAAGTGTCCCATGAGGCCACCGTTTCCAAACCATCAGATGAGGCAG	472
FEATURES	Location/Qualifiers	573	CCCTTCCAGTCT	632
source	1. .4516	473	AGCTTCCAGTCT	532
	/organism="Gallus gallus"	633	CCCCAGTCT	692
	/mol_type="mRNA"	533	GCTCCCGTCCAGCAACT	592
	/strain="CB"	693	TCCATGCT	746
	/db_xref="taxon:9031"	593	GCAGTACCCCATCT	652
	/clone="24a21"	747	TACCCCTCCGGGCTATTTCTACCACTTCCCACTTCCCACTTCCCACTTCCCACTTCCCACTTCCCA	806
	/cell_type="lymphocyte"	653	TATTCACAGGGTATTTCCAGCAGTCAATTTCCCACTTCCCACTTCCCACTTCCCACTTCCCACT	712
	/tissue_type="bursa"	807	AACCTCTAGTCT	860
	/clone_lib="rikeni"	713	AACCTCTAGTCT	772
	/dev_stage="2 weeks old"	861	CATCTCCCTCTCATACCCACTCTCCAAAGGCTCTCCAGTGTAGAGAGAGTCTCCCAAGAGGCTG	920
gene	1. .4516	773	CACAGGGTCACTACAGGGAGAAAAGTCACAAAGTATATCGACATCTCCGAGAGGAGC	832
5' UTR	<1..172	921	AAACATATGATGACACAGGACCCGAGACCACTAGTCTATGCGGAGGTGAGAGGCATCGG	980
CDS	/locus_tag="RCJMB04_24a21"	833	AAACACTAGCAGCAGCACCCGACACCGTGAGCACAGTCACTCCACGGTGTACAGGCATCGA	892
	/locus_tag="RCJMB04_24a21"	981	TC-----CCTGGATCGCGGAGCGAGCGGCGAGTCCCGACGAGAGAACAAGA-CAAGA-	1030
	/note="ORF1"	893	TCCAAATAACACCGGGGATCGTCCGGATCTCAGGCGCGGATCTCAGACAGGAGGAGCAGGAA	952
	/codon_start=1	1031	--CAGCGGTACAGATCTGATTTATGACCGGAGGAGAACACCATCTCGCCACCGCAGCTAC	1088
	/product="hypothetical protein"	953	AGCAGTCGCGCACCGGACTGAAATATGACAGAGAAAGAAATGTGCGCTCATCATAGAAGTTAT	1012
	/protein_id="CAG32380.1"	1089	GAAACGCGACAGAGCAGGAGAACCGGAGAGACACAGGCATCTCAGACAACCCGAAGATCACCA	1148
	/db_xref="GI:53134951"	1013	GAGCGTAGCAGGAGCGGTGACAGGGATAGGACATAGGCACCTGTACAGCAGAGAAGATCACCA	1072
	/db_xref="UniProt/TREMBL:O5ZIR3"	1149	TCTCTGGAAGGTCTTCAAAAAAGAGTATAGAGATCTGGAAGAGGATTACGGTTTATTCG	1208
	/translation="MSFHSGRCPGCGGPGGARTSAQTVPONLRQLHPQPSVOYV DQAAPATYSNPATSYMPREDFPVPP PVPFPFPAPPSPAPSPATATGQSTFPMPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFP YSTGYSQFPFPNYSQSNSSQSTNSGSHFRHTGQYQKESQSDRSPE RSKYDEHREHSHIHGRSHNDRDRGSPDRRQSSRHRTYDRGRMSPH HRSYRSRDRDRHRDRSPSPDRSYKDYKRAGSRSPSRERKPRWEEERW SSQSANKENYTSIKKSESENAKEDDELLKPVWVRC'THSESVYNDPMQV GSTVVGTGSKJLDKEFEELGKQAKAARPPWPPTKLDLDESSSECDSD DSSCSSSDSVDFVIAIKKKAHPDRHELMYNDPGQNDGDLCKCSAKARTG IRHGYPGSEPIKCPMTNNAAGLYHRTVSPPTNFLTDRPTVIEDHREYIPEF SMFSAPLTNIPLCKVIRENIDYTHIFIEEMMPENFCVKGLEFSYLPKDIILEYDW HLKGPSVENDSVCPHFMPFVFLPDGKGVLSMHOILLYLRLCNKALPVEEIA NMLQWEELEWQYAECKMIVTSPGKPSVRIOLDREQFPNDVITPPIIVHFGIR PAQLSYAGDPQYKMLKWSYVCKRLHANSPKVQADKQKLSQREALQIKRQKNTRR EYVLSLQSGFWKTIISDVCKQAHMLPVLTHIRYHQCLMHLDKLIGTYTFDRDCLLQ LAMTPHSHLNGPDRHARNLSNCGIRPKYKDRKHHMMKKGINTLINMSRL GODDPAPSRINNERLEFLDGAWEFLVTSVHLVYLPFTLEGGATYRTVAIVONHSLA MLAKKLEDRFMLYAGHPDLCHESDLRHAMANCEALIGAVYLGSLSEAKQLFGRLL FNDKLDRLVNLPLHPLQESNDRQLIETSPVQLKITEPEEALGVITFVRLIAR APTLRTGVNHLTLGNHQRMEFLGDSIMQVATEYLFTHFPDHHEGHLITLRSLLVN RTQAKVAELGQGFATITNDKRPVALTKTLADLLESFIAALYIDKOLEYVHTFMN VCFPRLKEFLINQWNPDSQLQOCCFLTREGKEPDIPLYKLTQTVGFSHARTYTV AVYFKERIGCGKGPSIQQAEMGAAMDALEKYNFPQWAHQKRFIERKYRELKEMRWE REHQRETDETEARK"	1268		
	4184..>4516	1117	-----CCTGGATCGGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1116
3' UTR	/locus_tag="RCJMB04_24a21"	1269	GATTCTTGGGCCCCACCCCTCGAGATTGTGAATCATCTCGTCCCAAGTAGGAGGAGGAGGAGG	1328
ORIGIN		1117	-----CCGCTCTCCAAGTAGAGAAAGAAAG	1141
Query Match	53.8%; Score 2564; DB 5; Length 4516;	1329	AGAGCTCGTTGGGAGGAGAAAAAGACCGTTGGAGTGCAACACAGAGTTCTTGGCAAGAC	1388
Best Local Similarity	76.7%; Pred. No. 0;			
Matches 3315; Conservative	0; Mismatches 870; Indels 139; Gaps 9;			

1142 AGACCCCGATGGAGGAGAAAGAGAAAGTGCTGTGAGAGCCAGAGTGCCAAAGAA 1201
1389 AAGAACTATACCTCAATCAAGGAAAGAGCCGAGGAGACCATGCTGACAAAGATGAG 1448
1202 AAAAACTACATAGTATCAAGACAAAGAGTCAGAGGAGAAATGCATCTGAAAAAATGAA 1261
1449 GAGGAAGAGAAAGAACTTCTTAAGCCTGTGTGATTCGATGCACTCATTTCAAGAAACTAC 1508
1262 GATGAAGATGATGAATTACTTAACACAGTCTGGGTTGCTGTACCCCAITTTCTGAAGCTAT 1321
1509 TACTCCAGTGAACCCCATGGATCAGTGGGAGATTTCTACAGTGGTGGAAAGAGTAGGCTT 1568
1322 TATTCAATGACCCCATGGATCAAGTGGGAGCTCTACTGTGTGGTGGACACAGCAAGCTC 1381
1569 CGTGACTTATATGACAAATTTGAGAGAGAGTTGGGGAGGAGGCAAGAAAGGCAAGCT 1628
1382 CGAGATCTGTATGAGAAAGTTGAGGAGGAGCTAGGAAAGACGACAGCAAGGCAAGCA 1441
1629 GCTCGGCTCCGTGGGAACCTCCAAAGACGAGCTCGATGAGATTTAGAGAGTCCAGT 1688
1442 GCCAGGCCACCATGGAGGCCGCGAAGACCAAGCTGGATGAAGATCTTGAGAGCTCCAGT 1501
1689 GAATCCGAGTGTGAGTCTGATGAGGACAGCACCTGTTCTTAGCAGCTTCAGACTCTGAACTT 1748
1502 GAGTCAGATGGACTCTGATGATGACAGCAGCTGCTCCAGCAGTTCAGATTCGGAAGTT 1561
1749 TTTGAGCTTATTTGACAAATCAAAACGCAAAAGGCCACCTGACCGCATTCATGATGAA 1808
1562 TTTGATGTCAATTCAGAAATTAAGCGTAAGAAAGCACACCTGATCGTCTTCATGAGGAA 1621
1809 CTTTGGTCAACGATCCAGCCAGATGATGATGAGACCACTCTGCAAAATGACGCGCAAG 1868
1622 CTGTGTAACAATGATCCAGACAGATGAATGATGAGCCGCTGTGCAAGTGACGCGCTAAA 1681
1869 GCAAGACGACAGGAATTAAGGACACAGATTTATCTTGGAGAGAGGCCATCAAGCCCTGT 1928
1682 GCCGACGACAGGATAGACATGCAATTAATCTTGGTGAAGAGCCCATTAACCATGT 1741
1929 CGTCTTATGACCAAAATGCTGGCAGACTTTTCCACTACCGGATCACAGTCTCCCGCT 1988
1742 CGCCCCATGACCAACAAATGCTGGAAGACTGTACCAATTAATCGAATTAACAGTGCACCTCC 1801
1989 ACGAACTTTTAACTGACAGGCCAACTGTTATGATAGATGATGATGATGATGATGATGAT 2048
1802 ACAAACTTCTTAACAGACAGACCTACTGTTATGATGATGATGATGATGATGATGATGAT 1861
2049 GAAGGATTTTCTATGTTGACATGCCCTGCCCAATATTCACCTGTGTAAAGTAAAT 2108
1862 GAAGGATTTCTCCATGTTTTCATGCCCCCACTGACAAACATTCCTCTATGTAAAGTAAATC 1921
2109 AGATTCACATAGACTACAGATTCATTTCAATGAGAGATGATGATGATGATGATGATGATGAT 2228
1922 AGATTTAAATGACTATACAAATTCATTTCAATGAGAGATGATGATGATGATGATGATGAT 1981
2169 GTGAAAGGCTTGAATCTTTTCACTGTTTCTTATTCAGAGATTTTGGAAATATATGAC 2228
1982 GTGAAAGGCTTGTAGCTGTTCTTCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2041
2229 TGGAAATTTTAAAGGCTTGTGTTGTAAGACAGCCCTCCCTGCTGCCCAAGATTTCAATTTTC 2288
2042 TGGCAATTTTAAAGGCTTGTGTTGTAAGATGATTTCTGTTTCTGCCCCAGATTTCACTTC 2101
2289 ATGCCAGCTTTGTAGATTTTCTTCAGATGAGGAAAGAGATGCTGCTCATGACCCAG 2348
2102 ATGCCCGATTTGTAGATTTTCTTCAGATGAGGAGGAGGAGTCTGTCATGATGATGATGAT 2161
2349 ATTCTCTCTACTTGTAAAGTGCAGCAAGCCCTGCTGCTGAGGAGGATTTGCCAAT 2408
2162 ATTCTTCTCTACTTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2221
2409 ATGCTTCAGTGGGAGGAGCTGGAGTGGCAGAAATATGCAAGAAATGCAAGGCAATGATGAT 2468

2222 ATGCTTTCAGTGGGAAGAGCTGGAATGGCAGAAATATGCAAGAGAAATGCAAAAGGATGATC 2281
2469 GTTACCAACCTTGGGAGCAGAAACCAAGCTCTGTCCGTATCGATCAACTGATCGTGAACAG 2528
2282 GTTACAGCCCTGGGATGAAACCCAGCTCAGTTCGTATGATCAACTGACCGTGAACAA 2341
2529 TTTCAACCCCGATGTGATTAATCTTTCCGATTATCGTCCACTTTGGGATACGCCCTGCACAG 2588
2342 TTTCAATCTGTATGATTAATCTTTCCCATTAATTTGCCATTTGGGATACGACTGCTCAA 2401
2589 TTTGAGTTATGCAAGGAGACCCACAGTACCAAAATCTGTGGAAGATTTATGTAAACTTCGC 2648
2402 CTCAGCTACGCTGGAGATCCTCAATACCAAAATATGGAAGATTTATGTGAAGCTGCT 2461
2649 CACCTCTAGCAAAATAGTCCGAAAGTCAAACTGACAAACAGAGCTGGCAGACAGG 2708
2462 CACCTGCTAGCTAAACAGTCCCAAGTCAAAAGCTGATAAAGAGAAATTAACACAGAGA 2521
2709 GAGGAAGCCCTCCAAAAATACGCAAGAAATACAATGAGACGAGAAATGAAACGCTGGAG 2768
2522 GAGGAAGCACTCGAAGAAATCGGCAAGAAACAGATGAGACGCTGAGTGTGCTGAG 2581
2769 CTAAGTAGCAAGGATTTGGAATCTGGCATCCGTCTGTGATGCTGTGTCAGCATGCAATG 2828
2582 TTTGAGTAGCCAGGATTTCTGGAAGACGCGGATACGCTCTGATGTCTGCCAGCAGCAATG 2641
2829 ATGCTACCTGTTCTGACCCATCATATCCGCTACCACTCAATGCTCAATGCTTGGACAG 2888
2642 ATGCTTCTGCTCTCACCCATATTCGCTACCATCAGTGTCTGATGATTTGGACAG 2701
2889 TTTGATAGGATATACCTTTCCAGATCGTCTGTGTGACGCTGGCCATGACTCATCCAAAT 2948
2702 TTTGATAGGCTACATTTTTCAGATAGTGTGCTGCTGACGTTGCCATGACTCATCCGAGC 2761
2949 CATCAATTTAAATTTTGGATGAATCTGATCATGCCAGGAATTCATTAATCTGTGGA 3008
2762 CATCACTTTAAATTTTGGATGAATCCAGACCATGCCAGAAATTCCTTATCCAACTGGG 2821
3009 ATTGGCAGCCCAATACGAGACAGAAAGTTTCATCATCATGCAATGCGGAGAGAGG 3068
2822 ATTGACAGCCAAAGTATGGGACAGAAAGTTTCCATATGCATATGAGAAAGG 2881
3069 ATTAAACACTTGCATAAATATCATGTCACGCTTGGCCAGATGACCCAACTCCCTCGAG 3128
2882 ATAAACACTGATTAATTAATTAATGTCCTGTTGGGACAGATGATCCAGCTCTTCCAGG 2941
3129 ATTAACCAAAATGAAACGGTTGGAATTTCTGGGTGATGCTGTTGTTGGAATTTCTGACAGC 3188
2942 ATTAACCAAAATGAGCGTTTAGAAATTTCTGGGAGATGCTGTGGTGGAGTTCTTTAA 3001
3189 GTCCATTTGTACTATTTGTTTCTTAGTCTGGAAGAGGAGGATTAACACCTATCCGACT 3248
3002 GTCCATTTGTACTATTTGTTTCTTAGTCTGGAAGAGGAGGATTAACACCTATCCGACT 3061
3249 GCATTTGTTTCAAGATCAGCACTTGCATGCTAGCAAAAGAACTTTGAACTGGATTCATTT 3308
3062 GCATAGTTCAAGACAGCATCTGSCCATGCTGGCTTAAAGCTGGAATAGATCGATTT 3121
3309 ATGCTGTATGCTACCGGCTGACCTTTGTAGAGAAATCGGACCTTCGACATGCAATGCAATGGCC 3368
3122 ATGCTTTATGCTCATGGTCCGACCTTTGCAAGGAAATCGGATCTCCGCCATGCCATGGCC 3181
3369 AATTGTTTGAAGCTTAATAGAGCTGTTTACTTTGGAGGAGGAGCTGAGAGAGCAAG 3428
3182 AACTGTTTGAAGCACTAATAGGACGATTAATGAGGGTGTGTTGAGAGAGCAAA 3241
3429 CAGTTATTTGAGCGCTTGCTCTTTAATGATCCGACCTGCGGAGTCTGGCTCAATTAAT 3488
3242 CAATTAATTTGAGCGTGTCTCTCAATGACAGGACCTGCGGAGTATGAGCTTAATTAAT 3301
3489 CCTCTCCACCCACTCAACTCAAGAGCCAAATATGATGACCACTTAATTAATTAATTAAT 3548
3302 CCACTACACCCACTCCAACTGCAGGAGTCTAATACTGACAGGCACTCATTTGAGACCTCT 3361


```
QY 3549 CAGTTTCTACAAAACCTTACTGAGTTTGAAGAAGCAATTGGAGTAATTTTACTCATGTT 3608
|||
DB 3362 CCAGTTCTTCAAAAGCTTACGGAGTTTGAAGAGCAATTGGAGTCATCTTTACTCATGTT 3421
|||
QY 3609 CGACTTCTGCAAGGCGCATTCACATTGAGAACTGTGGATTTAAACCATCTGACCCCTAGGC 3668
|||
DB 3422 CGACTTCTAGCAGCTGATCTCACTGTGAGGACAGTAGGCTTTAACCACTGACTCTAGGC 3481
|||
QY 3669 CACAATCAGAGAATGGAATTCCTAGTGACTCCATAATGCAACTGTGTAGCCACAGAGTAC 3728
|||
DB 3482 CACAATCAGAGGATGAAATTCCTGGGTGACTCCATAATGCAAGTGTGTAGCCACAGAGTAT 3541
|||
QY 3729 TTATTCATTTCATCCAGATCATCATGAAGGACACTTAACTTTGTGGAGCTCTTTG 3788
|||
DB 3542 TTATTCATCATATTTCCCGCATCATCATGAAGGCACTTAAACGCTGTGTGAAGCTCTCTG 3601
|||
QY 3789 GTGAATAATAGAACTCAGGCGAAGGTAGCGAGAGCTGGCATGACAGAGTACGCCATA 3848
|||
DB 3602 GTGAACACAGGACTCAGGCGCAAGGTGGCAGAGAGCTGGGAATGCAAGATTTGCCATC 3661
|||
QY 3849 ACCAACGACAGAACCAAGAGGCTGTGGCGCTTTCGACCAAGACCTTGGCGGACCTTTTG 3908
|||
DB 3662 ACTAATGACAGACCAAAAGACCGTAGCTCTTCGAACCTAAACTCTGGCAGATCTCTTG 3721
|||
QY 3909 GAATCATTTATTCAGCGCTGTACACTGATAGGATTTGGAATATGTTTCACTTTTCATG 3968
|||
DB 3722 GAATCCTTTATTCGTCCTGTACATGTGATAAGATTTGGAATATGTTTCACTTTTCATG 3781
|||
QY 3969 AATGCTGCTCTTTCCAGATTTGAAGAAATTCATTTTGAATCAGGATTTGAATGACCCC 4028
|||
DB 3782 AATGTATGCTTTTTTCCAGACTAAAGGAATTCATTTTAAATCAAGATTTGAATGACCT 3841
|||
QY 4029 AAATCCAGCTTCAGCAGTGTGCTTGACACTTAGCAGAGAGGAAAGAGCCAGACATT 4088
|||
DB 3842 AAATCTCAGCTTCAGCAGTGTGCTTGACTCTCAGGACAGAGGAAAGAGCCAGACATT 3901
|||
QY 4089 CTTCTGTAACAAGCTCTGCAGACAGTGGGCGCATCCCATGCGGAACTTACACTGTGGCT 4148
|||
DB 3902 CTTCTTTACAAGACTCTGCAGACAGTGGGACCATCTCATGCCAGGACGTACACTGTAGCT 3961
|||
QY 4149 GTTTATTTCAAGGAGAAAGATAGCTGTGGAAAGGACCAAGTATTCAGCAAGCGGAA 4208
|||
DB 3962 GTTTATTTCAAGGAGAAAGATAGGATGTGTAAGGCGCCCAAGTATTCACAAGAGCAAA 4021
|||
QY 4209 ATGGGAGCAGCAATGATGCGCTTGAAAAATATAATTTTCCCAAGATGCCCATCAGAAG 4268
|||
DB 4022 ATGGGAGCTGCAATGATGCACTTGAAATATATACTTTCCCAAGATGCCCATCAGAA 4081
|||
QY 4269 CGTTTCATCGAAGGAGTACAGACAAGAGTTAAAGAAATAGAGTGGGAAAGAGAGCAT 4328
|||
DB 4082 CGGTTCAATTGAACGGGAAGTACAGACAAGAGTTGAAAGAAATAGGTGGGAAAGAGAGCAT 4141
|||
QY 4329 CAAGAGAGAGCCAGATGAGACTGAGACATCAGAAAT-AAAGAGGGCATGCAAG-T 4386
|||
DB 4142 CAAGAGAGAGAAACAGATGAGACTGGAAGAGCAAGGAAATAAAGGAGGCGCGGAAGCA 4201
|||
QY 4387 GTGGAGTATTTACTTCTCAGTAACTGTGACTGTTGCTCTATTGAGACCTAGCTAGTTT 4446
|||
DB 4202 GTGGCANATTTTACTTAACTTAATCTGACTGTGCGCTTATGGAGACCTTAACCTAGTTT 4261
|||
QY 4447 CTTGCAGACAATGAACGAAGTGTGCTCATTTGAAATAAATAACAGAGTCA-AAATCGCTATT 4505
|||
DB 4262 -TTGCAGATGATGAATGAGAGTGCCTGTTGATATCGAATAGAAAATCATATCTTACT 4320
|||
QY 4506 GTTGTTTTAACTGATCTGTTTATAGCTGGATGCTTCTTATTAACAAGATATTAGATTTTCT 4565
|||
DB 4321 TTGAAGTGTGTATAGTGTCTTTTAGTTAGTTTGGTTTAAAGAAATCAGGTTTTTTTCTTCT 4380
|||
QY 4566 TCTA 4569
|||
DB 4381 GGTA 4384
```

```
RESULT 13
BC024261
LOCUS
DEFINITION
Homo sapiens ribonuclease III, nuclear, mRNA (cDNA clone
IMAGE:3631271).
ACCESSION
BC024261
VERSION
BC024261.1 GI:22028101
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini,
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2607)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klauser,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Udino,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2607)
NIH MGC Project
Direct Submission
Submitted (26-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: cgabbs-r@mail.nih.gov
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Martin Hirst, Thomas Zeng, Ryan Morin, Michelle Moksa, Johnson
Pang, Diana Mah, Jing Wang, Kieth Fichter, Eric Chuah, Allen
Delaney, Rob Kirkpatrick, Agnes Baross, Sarah Barber, Mabel
Brown-John, Steve S. Chand, William Chow, Ryan Babakaiff, Dave
Wong, Corey Matsuo, Jaclyn Beland, Susan Gibson, Luis delRio, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Mike R. Mayo, Josh Moran, Diana Palmquist, JR
Santos, Duane Smalhus, Jeff Stott, Miranda Tsai, George Yang,
Jacquie Schein, Asim Siddiqui, Steven Jones, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLML at: http://image.llnl.gov
Series: IRAL Plate: 40 Row: b Column: 20.
Location/Qualifiers
1..2607
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3631271"
/tissue_type="Uterus, endometrium adenocarcinoma"
```

```
FEATURES
source
```

/clone_lib="NIH_MGC_44" /lab_host="DH10B-r" /note="Vector: pOTB7"									
Query Match 53.0%; Score 2525.2; DB 8; Length 2607; Best Local Similarity 99.9%; Pred. No. 0; Matches 2527; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
QY	2126	CACGATTCATTTCAATTGAAGAGATGATGCGGAGAAATTTTGTGTGAAGGGCTTGAAC	2185						
Db	1	CACGATTCATTTCAATTGAAGAGATGATGCGGAGAAATTTTGTGTGAAGGGCTTGAAC	60						
QY	2186	CTTTTCACATGTTCCCTATTCAGAGATAATTTTGGAAATTTATATGACTGGAATCTTTAAAGGTCC	2245						
Db	61	CTTTTCACATGTTCCCTATTCAGAGATAATTTTGGAAATTTATATGACTGGAATCTTTAAAGGTCC	120						
QY	2246	TTTGTTTGAAGACAGCCCTCCCTGCTGCCCAAGATTTTCAATTTTCATGSCCAGCTTTTGTAAAG	2305						
Db	121	TTTGTTTGAAGACAGCCCTCCCTGCTGCCCAAGATTTTCAATTTTCATGSCCAGCTTTTGTAAAG	180						
QY	2306	ATTCTTCCAGATGGAGGAAAGAAAGTGTCTGTCATGCAACGATTTCTCTGTACTTGT	2365						
Db	181	ATTCTTCCAGATGGAGGAAAGAAAGTGTCTGTCATGCAACGATTTCTCTGTACTTGT	240						
QY	2366	AAAGTCACGAAGCCCTGGTGCCTGAGGAGGATTCCTATATGCTTCAGTGGGAGGA	2425						
Db	241	AAAGTCACGAAGCCCTGGTGCCTGAGGAGGATTCCTATATGCTTCAGTGGGAGGA	300						
QY	2426	GCTGGAGTGGCAGAAATATGCAAGAAATGCAAAAGGCATGATTTGTACCAACCCCTGGGAC	2485						
Db	301	GCTGGAGTGGCAGAAATATGCAAGAAATGCAAAAGGCATGATTTGTACCAACCCCTGGGAC	360						
QY	2486	GAAACCAAGCTGTGCTGATCGATCAACTGGAATCGTGAACAGTTTCAACCCCGATGTGAT	2545						
Db	361	GAAACCAAGCTGTGCTGATCGATCAACTGGAATCGTGAACAGTTTCAACCCCGATGTGAT	420						
QY	2546	TACTTTTCGATATCGTCCATCTTTGGGATACGCCCTGCACAGTTGAGTTATGCAAGGGA	2605						
Db	421	TACTTTTCGATATCGTCCATCTTTGGGATACGCCCTGCACAGTTGAGTTATGCAAGGGA	480						
QY	2606	CCACAGTACCAAAAATCTGTGAAGAGTTATGTGAAACTTTGSCCACTCTCTAGCAAAATAG	2665						
Db	481	CCACAGTACCAAAAATCTGTGAGAGTTATGTGAAACTTTGSCCACTCTCTAGCAAAATAG	540						
QY	2666	TCCAAAGTCAACAAACTGA CAAACAGAGCTGGCAGCAGAGGGAGAAAGCCCTCCAAA	2725						
Db	541	TCCAAAGTCAACAAACTGA CAAACAGAGCTGGCAGCAGAGGGAGAAAGCCCTCCAAA	600						
QY	2726	AATACGGCAGAGAAATACATGAGACGAGAGTAAACGGTGGAGCTAAAGTAGCCAAAGATT	2785						
Db	601	AATACGGCAGAGAAATACATGAGACGAGAGTAAACGGTGGAGCTAAAGTAGCCAAAGATT	660						
QY	2786	CTGGAAAACTGGCATCCGTTCTGATCTGTGATGATGCAATGATGCAATGATGCTACCTGTTCTGAC	2845						
Db	661	CTGGAAAACTGGCATCCGTTCTGATCTGTGATGATGCAATGATGCAATGATGCTACCTGTTCTGAC	720						
QY	2846	CCATCATATCCGTTACCA CCAATGCTTAATGCAATTTTGGCAAGTTGATAGGATATATCTTT	2905						
Db	721	CCATCATATCCGTTACCA CCAATGCTTAATGCAATTTTGGCAAGTTGATAGGATATATCTTT	780						
QY	2906	CCAGATCGTTGTCTGTGGAGCTGGCCATGACTCATCCAGTCATCTTAAATTTTGG	2965						
Db	781	CCAGATCGTTGTCTGTGGAGCTGGCCATGACTCATCCAGTCATCTTAAATTTTGG	840						
QY	2966	AATGAATCTTGATCATGCCAGGAATTCATTTATCTAACTGTGGAAATTCGGCAGCCCAATA	3025						
Db	841	AATGAATCTTGATCATGCCAGGAATTCATTTATCTAACTGTGGAAATTCGGCAGCCCAATA	900						
QY	3026	CGGAGACAGAAAAAGTTTCATCACATGCACATGCGGAAAGAAAGGATTTAA CACTTGTATAA	3085						
Db	901	CGGAGACAGAAAAAGTTTCATCACATGCACATGCGGAAAGAAAGGATTTAA CACTTGTATAA	960						
ORIGIN									
QY	3086	TATCATGTCAAGCCTTGGCCAAAGATGACCCAACTCCCTCGAGGATTAACCAACAATGAACG	3145						
Db	961	TATCATGTCAAGCCTTGGCCAAAGATGACCCAACTCCCTCGAGGATTAACCAACAATGAACG	1020						
QY	3146	GTTTGAAATTCCTGGGTGATGCTGTTGTTGAAATTTCTGACCAGCGTCCATTGTACTATTTT	3205						
Db	1021	GTTTGAAATTCCTGGGTGATGCTGTTGTTGAAATTTCTGACCAGCGTCCATTGTACTATTTT	1080						
QY	3206	GTTTCTCTAGTCTGGAAGAGGAGGATTAGCAACCTTATCGGACTGCAATTTGTTTCAAGATCA	3265						
Db	1081	GTTTCTCTAGTCTGGAAGAGGAGGATTAGCAACCTTATCGGACTGCAATTTGTTTCAAGATCA	1140						
QY	3266	GCACCTTGCATGCTAGCAAAAGAAACCTTGAACCTGGATCCATTTATGCTGTATGCTCACGG	3325						
Db	1141	GCACCTTGCATGCTAGCAAAAGAAACCTTGAACCTGGATCCATTTATGCTGTATGCTCACGG	1200						
QY	3326	GCCTGACCTTTGTAGAGAAATCGGACCTTGCACATGCAATGGCCAAATTTGTTTGAAGCGTT	3385						
Db	1201	GCCTGACCTTTGTAGAGAAATCGGACCTTGCACATGCAATGGCCAAATTTGTTTGAAGCGTT	1260						
QY	3386	AATAGGAGCTGTTTACTTTGGAGGAAAGCTGAGGAAGCCAAAGCAGTATATTTGGAGCGTT	3445						
Db	1261	AATAGGAGCTGTTTACTTTGGAGGAAAGCTGAGGAAGCCAAAGCAGTATATTTGGAGCGTT	1320						
QY	3446	GCTCTTTAATGATCGGACCTGCGGAAAGTCTGGCTCAATTTATCTCTCCACCCACTCCA	3505						
Db	1321	GCTCTTTAATGATCGGACCTGCGGAAAGTCTGGCTCAATTTATCTCTCCACCCACTCCA	1380						
QY	3506	ACTACAAGAGCCAAATATCTGATCGACAACTTATGAAACTTCTCCAGTTCTACAAAACCT	3565						
Db	1381	ACTACAAGAGCCAAATATCTGATCGACAACTTATGAAACTTCTCCAGTTCTACAAAACCT	1440						
QY	3566	TACTGAGTTTGAAGAGCAATTTGGAGTAAATTTTACTCATGTTTCGACTTCTCGAAGGCG	3625						
Db	1441	TACTGAGTTTGAAGAGCAATTTGGAGTAAATTTTACTCATGTTTCGACTTCTCGAAGGCG	1500						
QY	3626	ATTTCATTTGAAACTGTGGGATTTAAACCATCTGACCCCTAGGCCCAATACAGAGATGGA	3685						
Db	1501	ATTTCATTTGAAACTGTGGGATTTAAACCATCTGACCCCTAGGCCCAATACAGAGATGGA	1560						
QY	3686	ATTCTTAGGTGACTCCATATGCAACTGGTAGCCACAGAGTACTTATTCATTTCTTCCC	3745						
Db	1561	ATTCTTAGGTGACTCCATATGCAACTGGTAGCCACAGAGTACTTATTCATTTCTTCCC	1620						
QY	3746	AGATCATCATGAAGGACACTTAACTTTTGTGCGAAGCTCTTTGGTGAATAATAGAACTCA	3805						
Db	1621	AGATCATCATGAAGGACACTTAACTTTTGTGCGAAGCTCTTTGGTGAATAATAGAACTCA	1680						
QY	3806	GGCCAAAGGTAGCGGAGGAGCTGGGCATGCAAGAGTACGCCATAACCAACGACAAAGACCA	3865						
Db	1681	GGCCAAAGGTAGCGGAGGAGCTGGGCATGCAAGAGTACGCCATAACCAACGACAAAGACCA	1740						
QY	3866	GAGGCTGTGGCGCTTCGACCAAGACCTTGGCGGACCTTTTGGATCATTTATTCGACG	3925						
Db	1741	GAGGCTGTGGCGCTTCGACCAAGACCTTGGCGGACCTTTTGGATCATTTATTCGACG	1800						
QY	3926	GCTGTACACTGATAGGATTTTGGATATGTCATCTTTTCATGAATGTCGTCTCTTTCC	3985						
Db	1801	GCTGTACACTGATAGGATTTTGGATATGTCATCTTTTCATGAATGTCGTCTCTTTCC	1860						
QY	3986	ACGATTTGAAAGAAATTCATTTTGAATCAGGATTTGGATGACCCCAATCCAGCTTCAGCA	4045						
Db	1861	ACGATTTGAAAGAAATTCATTTTGAATCAGGATTTGGATGACCCCAATCCAGCTTCAGCA	1920						
QY	4046	GTGTGCTTTGACATTTAGGACAGAGGAAAGAGCCAGACATTCCTCTGTACAGACTCT	4105						
Db	1921	GTGTGCTTTGACATTTAGGACAGAGGAAAGAGCCAGACATTCCTCTGTACAGACTCT	1980						
QY	4106	GCAGACAGTGGGCGCCATCCCATGCCGACCTTACCTGTGGCTGTTTATTTCAAGGGAGA	4165						
Db	1981	GCAGACAGTGGGCGCCATCCCATGCCGACCTTACCTGTGGCTGTTTATTTCAAGGGAGA	2040						
QY	4166	AAGAATAGGCTGTGGGAAAGGACCAAGTATTTCAGCAAGCGGAAAATGGGAGCAGCAATGGA	4225						

Dd	121	AGGAAAGGAAGTGTCTCCATGCACACAGATTCTCCTGTACTTTGTTAAGGTGCAGCAAGC	180	Qy	3461	GGACCTGCGCGAAGTCTGGCTCAATTATCTCTCCACCACCTCCAACTCAAGAGCCAAA	3520
Qy	2381	CCTGGTGCCTGAGGAGGATTTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGCAGAA	2440	Dd	1261	GGACCTGCGCGAAGTCTGGCTCAATTATCTCTCCACCACCTCCAACTCAAGAGCCAAA	1320
Dd	181	CCTGGTGCCTGAGGAGGAGATTTGCCAATATGCTTCAGTGGGAGGAGCTGGAGTGCAGAA	240	Qy	3521	TACTGATCGACAACTTATTTGAAACTTTCCAGTTCTACAAAACTTACTGTGAGTTGAAGA	3580
Qy	2441	ATATGCGAAGAAATGCAAGAGGATGATTTGTACCAACCCCTGGGACGAAACCAAGCTCTGT	2500	Dd	1321	TACTGATCGACAACTTATTTGAAACTTTCCGTTCTACAAAACTTACTGTGAGTTGAAGA	1380
Dd	241	ATATGCGAAGAAATGCAAGAGGATGATTTGTACCAACCCCTGGGACGAAACCAAGCTCTGT	300	Qy	3581	AGCAATTGGAGTAAATTTTACTCATGTTTCGACTTCTGGCAAGGGCATTCACATTTCGAAAC	3640
Qy	2501	CGGTATCGATCAACTGGATCGTGAACAGTTCAACCCCGATGTGATTTTCCGATTAT	2560	Dd	1381	AGCAATTGGAGTAAATTTTACTCATGTTTCGACTTCTGGCAAGGGCATTCACATTTCGAAAC	1440
Dd	301	CGGTATCGATCAACTGGATCGTGAACAGTTCAACCCCGATGTGATTTTCCGATTAT	360	Qy	3641	TGTGGGATTTAACTCATCTGACCTAGGCCACCAATCAGAGAAATGGAATTCCTAGGTGACTC	3700
Qy	2561	CGTCCACTTTGGGATAGCCCTGCGACAGTTGAGTTATGCAAGAGACCCACAGTACCAAAA	2620	Dd	1441	TGTGGGATTTAACTCATCTGACCTAGGCCACCAATCAGAGAAATGGAATTCCTAGGTGACTC	1500
Dd	361	CGTCCACTTTGGGATAGCCCTGCGACAGTTGAGTTATGCAAGAGACCCACAGTACCAAAA	420	Qy	3701	CATAATGCAACTGTGTAGCCACAGAGTACTTATTTCAATTCATTTCCAGATCATCATGAAGG	3760
Qy	2621	ACTGTGGAAAGATTATGTGAAACTTTCGCCACCTCTAGCAAAATAGTTCCTCAAAAGTCAAA	2680	Dd	1501	CATAATGCAACTGTGTAGCCACAGAGTACTTATTTCAATTCATTTCCAGATCATCATGAAGG	1560
Dd	421	ACTGTGGAAAGATTATGTGAAACTTTCGCCACCTCTAGCAAAATAGTTCCTCAAAAGTCAAA	480	Qy	3761	ACACTTAACTTTGTTGGGAAAGCTCTTTGGTGAATATAGAACTCAGGCCCAAGGTAGCGGA	3820
Qy	2681	AACTGACAAACAGAGCTGGCACAGAGGGAGGAGCCCTCCAAAAAATACGGCAGAGAA	2740	Dd	1561	ACACTTAACTTTGTTGGGAAAGCTCTTTGGTGAATATAGAACTCAGGCCCAAGGTAGCGGA	1620
Dd	481	AACTGACAAACAGAGCTGGCACAGAGGGAGGAGCCCTCCAAAAAATACGGCAGAGAA	540	Qy	3821	GGAGCTGGGCATGCAAGGAGTACGCCATAACCAACGACCAAGACCAAGAGCCCTGTGGCGCT	3880
Qy	2741	TACATGAGACAGAGAAATTAACGGTGGAGCTAAGTAGGCAAGGATTTCTGAAAACTGGCAT	2800	Dd	1621	GGAGCTGGGCATGCAAGGAGTACGCCATAACCAACGACCAAGACCAAGAGCCCTGTGGCGCT	1680
Dd	541	TACAAATGAGACAGAGAAATTAACGGTGGAGCTAAGTAGGCAAGGATTTCTGAAAACTGGCAT	600	Qy	3881	TCGCAACCAAGACCTTTGGCGACCTTTTGGAAATCATTTATTTGACGCGTGTAACATGATAA	3940
Qy	2801	CGTTCCTGATGCTGTGACGATGCAATGATGCTACCTGTCTTGACCCATCATATCGCTA	2860	Dd	1681	TCGCAACCAAGACCTTTGGCGACCTTTTGGAAATCATTTATTTGACGCGTGTAACATGATAA	1740
Dd	601	CGTTCCTGATGCTGTGACGATGCAATGATGCTACCTGTCTTGACCCATCATATCGCTA	660	Qy	3941	GGATTTGGAATATGTTTATCATCTTTTATGAAATGCTCTCTTTCCACGATTCGAAAGAAAT	4000
Qy	2861	CCACCAATGCCATTAATGCAATTTGGACAAAGTTGATAGGATATCTTTCCAAAGTCTGTGCT	2920	Dd	1741	GGATTTGGAATATGTTTATCATCTTTTATGAAATGCTCTCTTTCCACGATTCGAAAGAAAT	1800
Dd	661	CCACCAATGCCATTAATGCAATTTGGACAAAGTTGATAGGATATCTTTCCAAAGTCTGTGCT	720	Qy	4001	CATTTTGAATCAGGATTTGGAATGACCCCAATTCACAGCTTCAGCAGTGTGTCTTGACACT	4060
Qy	2921	GTTCAGCTGGCGATGACTCATCCAGTTCATCATTTAAATTTTGGNAATGAAATCTGTATCA	2980	Dd	1801	CATTTTGAATCAGGATTTGGAATGACCCCAATTCACAGCTTCAGCAGTGTGTCTTGACACT	1860
Dd	721	GTTCAGCTGGCGATGACTCATCCAGTTCATCATTTAAATTTTGGNAATGAAATCTGTATCA	780	Qy	4061	TAGGACAGAGGAAAGAGCCAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCC	4120
Qy	2981	TGCCAGGAATTCATTAATCTAACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAGT	3040	Dd	1861	TAGGACAGAGGAAAGAGCCAGACATTCCTCTGTACAAGACTCTGACAGAGTGGGCC	1920
Dd	781	TGCCAGGAATTCATTAATCTAACTGTGGAATTCGGCAGCCCAATACGGAGACAGAAAAGT	840	Qy	4121	ATCCCATGCCCGAACCTTACACTGTGGCTGTATTTTCAAGSGGAGAAATAGGCTGTGG	4180
Qy	3041	TCATCACATGCATGCGGAAGAAAGGGATTAACACTTTGATAAAATATCATGTCAAGCT	3100	Dd	1921	ATCCCATGCCCGAACCTTACACTGTGGCTGTATTTTCAAGSGGAGAAATAGGCTGTGG	1980
Dd	841	TCATCACATGCATGCGGAAGAAAGGGATTAACACTTTGATAAAATATCATGTCAAGCT	900	Qy	4181	GAAAGACCAAGTATTTCAGCAAGCGGAAATGGAGCAGCAATGGGATGCGCTTGAAAAATA	4240
Qy	3101	TGCCCAAGATGACCCAACTCCCTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGG	3160	Dd	1981	GAAAGACCAAGTATTTCAGCAAGCGGAAATGGAGCAGCAATGGGATGCGCTTGAAAAATA	2040
Dd	901	TGCCCAAGATGACCCAACTCCCTCGAGGATTAACCAATGAACGGTTGGAAATTCCTGGG	960	Qy	4241	TAAATTTCCCGCAGATGGCCCATCAGAAGCGGTTTCATTCGAACGGAAATCAGACAAGAGTT	4300
Qy	3161	TGATGCTGTGTGAAATTTCTGACCAAGCGTTCATTTTGTACTATTTGTTTCCCTAGTCTGA	3220	Dd	2041	TAAATTTCCCGCAGATGGCCCATCAGAAGCGGTTTCATTCGAACGGAAATCAGACAAGAGTT	2100
Dd	961	TGATGCTGTGTGAAATTTCTGACCAAGCGTTCATTTTGTACTATTTGTTTCCCTAGTCTGA	1020	Qy	4301	AAAAGAAATGAGGTGGGAAAAGAGCATCAAGAGAGAGCCAGATGAGACTGGAAGACAT	4360
Qy	3221	AGAAGAGGATTAAGCAACCTATCGGACTGCAATTTTTCAGAATTCAGACCTTTGCCATGCT	3280	Dd	2101	AAAAGAAATGAGGTGGGAAAAGAGCATCAAGAGAGAGCCAGATGAGACTGGAAGACAT	2160
Dd	1021	AGNAGAGGATTAAGCAACCTATCGGACTGCCATTTTCAGAATTCAGACCTTTGCCATGCT	1080	Qy	4361	CAAGAAATAAGAGGAGGCAATGCAAGTGTGGAGTATTTTACTGCTCAGTAACTGTGACTGT	4420
Qy	3281	AGCAAGAAACTTTGAACTGGAATTCGATGATGCTGATGCTACGGGCTGACCTTTGTAG	3340	Dd	2161	CAAGAAATAAGAGGAGGCAATGCAAGTGTGGAGTATTTTACTGCTCAGTAACTGTGACTGT	2220
Dd	1081	AGCAAGAAACTTTGAACTGGAATTCGATGATGCTGATGCTACGGGCTGACCTTTGTAG	1140	Qy	4421	TGTCTATTGAGACCTTAGCCCTAGTTTCTGACAGCAATGAAAGAGTGTGCTCATTTGAAA	4480
Qy	3341	AGAACTCGGACCTTCGACATGCAATGCGCAATTTGTTTGAAGCGTTAATAGGAGCTGTTTA	3400	Dd	2221	TGTCTATTGAGACCTTAGCCCTAGTTTCTGACAGCAATGAAAGAGTGTGCTCATTTGAAA	2280
Dd	1141	AGAACTCGGACCTTCGACATGCAATGCGCAATTTGTTTGAAGCGTTAATAGGAGCTGTTTA	1200	Qy	4481	TAAATACAGAGTCAAAATCGCTATTTGTTTGAATGATCTGTTTTTACGTGATGCTGCT	4540
Qy	3401	CTTTGAGGGAAGCCTCGAGAGGAGCCCAAGCAGTTATTTGGACGCTTCTCTTTAATGATCC	3460	Dd	2281	TAAATACAGAGTCAAAATCGCTATTTGTTTGAATGATCTGTTTTTACGTGATGCTGCT	2340
Dd	1201	CTTTGAGGGAAGCCTCGAGAGGAGCCCAAGCAGTTATTTGGACGCTTCTCTTTAATGATCC	1260				

```

QY 4541 TTATTACAAAGTATTAGATTTTCTTCTATTATTAACGGAACCTTGACTTTGGTGAATGTG 4600
Db 2341 TTATTACAAAGTATTAGATTTTCTTCTATTATTAACGGAACCTTGACTTTGGTGAATGTG 2400

QY 4601 CATTACTCTCTTTTATTTTCTCTTTTAAATAATAAATAATTCAAGAAGCATA 4650
Db 2401 CATTACTCTCTTTTATTTTCTCTTTTAAATAATAAATAATTCAAGAAGCATA 2450

RESULT 15
BC057687
LOCUS
DEFINITION Mus musculus ribonuclease III, nuclear, mRNA (cDNA clone
IMAGE:3995734), partial cds.
ACCESSION BC057687
VERSION
KEYWORDS
SOURCE BC057687.1 GI:34784427
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2971)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2971)
NIH MGC Project
Direct Submission
Submitted (03-SEP-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 120 Row: g Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers
1. .2971

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:3995734"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMV."
/clone_lib="NCI_CGAP_Lu29"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
<1. .2971
/gene="Rnasen"
/feature="synonym: Rn3"
/db_xref="GeneID:14000"
/db_xref="MGI:1261425"
<1. .2830
/gene="Rnasen"
/codon_start=2
/product="Rnasen protein"
/protein_id="AAH5687.1"
/db_xref="GI:34784428"
/db_xref="GeneID:14000"
/db_xref="MGI:1261425"
/translation="STVVGTSRLDLYDKFEELGNROEKAKAARPPWPPKTKLDED
LESSSECEUDDSTSCSSSEVDFVIAEIKRKAHPDRDLHDLWYNDPQMDGDP
LCKSAKARRTGIRHSIYPGEBAIKPCRPMTNAGRLPHYRITVSPNTFLDRPTVI
EYDDIEYFEGFSMAHAPLNTPLCKVIRFNIDYTHFIEEMPFNFCVKGLELPSL
FLFRDILELYDNWLGKPLFDPSPCPRFHMFVRFLPDGKGKVLHQILLYLR
CSKALPREEIANMLQWELBWKVAECKGMIVNPGTPSSVRIDQDRQFQNEV
ITFPIIVHGIIPAQLSVAGDPQYOKLWKSIVKLRLHLLANSKVKOTDKQKLAOREA
LQIKQKNTMRREVTVELSSQSFWKIGRSDVCQHMLPLVTHIRYHQCLMHLDKL
IGYTFQDRCLQLAMTHPSHLGNPDHARNSISNGIRQPKYGRDKRVHMHMKK
GINTLINTMSRLQDDPTPSRINHRERFLDGVAVFELTSVHLVYLPFSLBEGGLAT
YRTAIVQHOHLAMAKLELDRLFMYAHGPDLCRESDLAHAMANGFEALIGAVYLEGS
LBEAKOLGRILLFNDPLREWNLVPLHPLQOEPTDRQLETSPVOKLTFEPEAI
GVIFTHRLARAFTLRTVGNHILGHNRMEFLGDSIMOLVATEYLFIHPDHHEG
HLTLRSLVNNRTQAKVABELGMQERYAITNDKTRPVALRTKTLADLESFIALYI
DKDLRYVHTFMNVCFPRKLKFLNODMNDPKSQCCCLTLRTGEGEPDIPLYKTLQ
TVGPSHARTYTVAVTFKGERIGCGKGPSIQAEAGMAAMDALEKYNFPQMAHOKRFIER
KYRQLKEMRWEREHQEREPEAEIKK"

```

ORIGIN

```

Query Match 50.9%; Score 2425.6; DB 9; Length 2971;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 2629; Conservative 0; Mismatches 319; Indels 2; Gaps 1;

QY 1542 TCTACAGTGGTTGGAAACGAGTAGGCTTCGTGACTTATATGACAAATTTTCAGGAGGAGTTG 1601
Db 2 TCGACTGTGTAGGACAAAGCGAGCTCCGTGATTTGTATGACAAATTTTCAGGAGGAGTTG 61

QY 1602 GGGAGCAGCAAGAAAAGGCCAAAGCTGCTCGGCGCTTCGTGGGAACTTCCAAAGACGAAG 1661
Db 62 GGGATAGGCAAGAGAGGCGCAAGCTGCCGCGCTTCGTGGGAGGCTCCGAAGACAAAG 121

QY 1662 CTCGATGAAGATTTAGAGATTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGAGCAGCACC 1721
Db 122 CTGGATGAAGATTTAGAGATTTCCAGTGAATCCGAGTGTGAGTCTGATGAGGAGCAGCACC 181

QY 1722 TGTTCTACAGCTCAGACTCTGAAGTTTTTGAAGTTTATGACAGAAATCAACGCAAAAG 1781
Db 182 TGTTCTACAGCTCAGACTCTGAAGTTTTTGAAGTTTATGACAGAAATCAACGCAAAAG 241

QY 1782 GCGCCACCTTGACCGACTTCATGATGAACCTTTGGTACACGATCCAGGCGCAGATGATGAT 1841
Db 242 GCTACCTTGACCGCTTCATGATGAACCTTCGTACACGATCCAGGCGCAGATGATGAT 301

QY 1842 GGACCACTCTGCAATGAGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1901
Db 302 GGACCGCTTTGCAATGAGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 361

QY 1902 CTTGAGAGAGAGGCGCATCAAGCCCTGCTGCTTATGACCAACAATGCTGCGCAGACTTTC 1961

```

FEATURES
source

Db 362 CCGGAGAGAGGCAATCAAGCCCTGCGTCCGATGACCAACAAACGCTGCGCGCTTTTC 421
QY 1962 CACTACCGGATCAGCTCTCCCGCCTACGAACTTTTAACTGCACAGGCCAACTGTTATA 2021
Db 422 CACTATCGGATCACCGTCTCCCGCCTACCACTTCTTAACGTACAGGCCAACAGTCATA 481
QY 2022 GAATACGATGATCACGAGTATATCTTTTGAAGGATTTTCTATGTTTGCACATGCCCCCTG 2081
Db 482 GAATACGATGATCACGAGTATATTTTGAAGGATTTTCTATGTTTGCACATGCTCCTCTG 541
QY 2082 ACCAATATTCACCTGTGTAAAGTAATTAATCAACATAGATACATGATCAATTTCAAT 2141
Db 542 ACCAATATTCACCTGTGTAAAGTAATTCGATTCACATAGATACATGATTTTATC 601
QY 2142 GAAGAGATGATGCGGAGATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTTCTTA 2201
Db 602 GAAGAGATGATGCGGAGATTTTGTGTGAAGGCTTGAACCTTTTCACTGTTTCTTA 661
QY 2202 TTCCAGAGATATTTTGAATTTATATGACTGGAATCTTTAAAGGTCCTTTTGTGAAGACAGC 2261
Db 662 TTCCAGAGATATTTTGAATTTATATGACTGGAATCTTTAAAGGTCCTTTTGTGAAGACAGC 721
QY 2262 CTTCTGCTGCTGCCAAGATTTCAATTTTCATGCGCAAGTTTGTAAAGATTTCTTCCAGATGA 2321
Db 722 CTTCTGCTGCTGCCAAGATTTCAATTTTCATGCGCAAGTTTGTAAAGATTTCTTCCAGATGA 781
QY 2322 GGAAGGAGATGCTGTCATGACACAGATTTCTCTGACTGTTGTTAAAGTGCACGAAGCC 2381
Db 782 GGAAGGAGATGCTGTCATGACACAGATTTCTCTGACTGTTGTTAAAGTGCACGAAGCC 841
QY 2382 CTGGTCCCTGAGGAGAGATTTGCCAATATGCTTCAGTGGGAGAGCTGGAGTGGCAGAAA 2441
Db 842 CTGGTCCCGAGGAGAGATTTGCCAATATGCTTCAGTGGGAGAGCTGGAGTGGCAGAAA 901
QY 2442 TATGCAAGAGAAATGCAAGGCAATGATTTGTTACAACTTGGGACGAAACCAAGCTCTGTC 2501
Db 902 TACGACAGAGAGTGCAGAGCAATGATGCTTTACCACTTGGGACGAAACCGAGCTCTGTC 961
QY 2502 CGTATCGATCACTGATCGTGAACAGTTTCAACCCCGATGATTAATTTTCCGATATC 2561
Db 962 CGCAATGATCAGCTGGATCGTGAACAGTTTCAACCCCGATGATTAATTTTCCGATATC 1021
QY 2562 GTCCACTTTGGGATACGCCCTGCACAGTTTGAGTTATGACGAGGACCCACAGTACCAAAA 2621
Db 1022 GTTCACTTTGGGATCCGCCCGCACAGTTTGAGTTATGACGAGGACCCACAGTATCAGAAA 1081
QY 2622 CTGTGGAAGATTTATGTGAACCTTGCACCTCTAGCAAAATAGTCCCAAGTCAAAACAA 2681
Db 1082 TTATGAAAAGCTTATGTGAACCTTGCACCTCTAGCAAAACAGTCCCAAGTCAAAACAA 1141
QY 2682 ACTGACAAACAGNAGCTGGCACAGGGGAGGAGCCCTCAAAAATACGGCAGAGAT 2741
Db 1142 ACTGACAAAGCAAGCTGGCTCAGAGGAGGAGCGCTCCAGAAAATACGGCAGAGAAC 1201
QY 2742 ACAATCAGACGAGAAGTTAAGGTGAGCTAAGTAGCAAGGATTCGGAATACTGSCATC 2801
Db 1202 ACATCGCGGAGAGTACCGGTGAGAGCTGAGTAGCCAGGATTCGGAATACTGSCATC 1261
QY 2802 CGTTCGTATGTCGTGACGATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTAC 2861
Db 1262 CGTTCGTATGTCGTGACGATGCAATGATGCTACCTGTTCTGACCCCATCATATCCGCTAC 1321
QY 2862 CACCAATGCTTAATGCAATTTGGACAAGTTGATAGGATATATTTTCAAGATGTTGTCG 2921
Db 1322 CATCAATGCTTAATGCAACCTGGACAAGTTGATAGGATATATTTTCAAGACCGTTGTCG 1381
QY 2922 TTGCACTGCGCATGACTCATCCAAGTCATCATTTAAATTTTGGATGATCTCTGATCAT 2981
Db 1382 CTCACGCTTGCATGACTCATCCGAGTCACCATTTAAATTTTGGATGATCTCTGACCAT 1441
QY 2982 GCCAGGAATTCATTTACTTAACCTGTGGAATTTGGCAGCCCAAAATACGGAGACAGAAAAGTT 3041
Db 1442 GCCAGGAATTCATTTACTGTGGAATTTGCAACACCCCAATATGAGACAGAGAAAAGTT 1501

QY 3042 CATCATCATGATCGGAGAGAAAGGATTAACACCTTGATAAATATCATGTACGCCCTT 3101
Db 1502 CATCATCATGATCGGAGAGAAAGGATTAACACCTTGATAAATATCATGTACGCCCTT 1561
QY 3102 GGCAGAGATGACCCAACTCCCTCGAGGATTAACCAATGAACGGTTGGAATTCCTGGGT 3161
Db 1562 GGCAGAGATGATCCAACTCCCTCAAGGATTAATCAACAGAAAGTTGAGATTCCTGGGG 1621
QY 3162 GATGCTGTTGTTGAATTTCTGACCGAGCTCCATTTGTACTATTTGTTTCTTAGTCTGGAA 3221
Db 1622 GATGCTGTTGTTGAGTTTCTGACCGAGTCCACCTGTACTCTGTTCTTCTAGCTTGGAG 1681
QY 3222 GAAGGAGATTAAGCAACCTATCGGATGCTGCAATTTGTTTCAAGTACAGCACCTTGCCATGCTA 3281
Db 1682 GAAGGGGCTTGGGACCTATCGGACCGCAATGTTTCAAGTACAGCACCTTGCCATGCTT 1741
QY 3282 GCAAGAAACTTTGAACTGGATCCATTTATGCTGTATGCTCACGGGCTGACCTTTGTAGA 3341
Db 1742 GCAAGAAACTTTGAACTGGATCCATTTATGCTGTATGCCCATGGGCCGACCTATGTAGA 1801
QY 3342 GAATCGGACCTTGCACATGCAATGGCCAAATGTTTGAAGCGTTAATAGAGCTGTTTAC 3401
Db 1802 GAATCAGATCTCCGCCATGCGATGGCCAAATGTTTGAAGCGCTTGATAGGAGCTGTTTAC 1861
QY 3402 TTGGAGGGAAGCTCGAGGAAGCCAAAGCAGTTTATTTGAGACGCTTGTCTTTTAAATGATCG 3461
Db 1862 TTGGAGGGAAGCTCGAGGAAGCCAAAGCAGTTTATTTGAGACGCTTGTCTTTTAAATGATCG 1921
QY 3462 GACCTGCGGAACTGCTGCTCAATTTATCTCTCCACCACTCAACTCAAGAGCCAAAT 3521
Db 1922 GACCTGCGGAACTGCTGCTCAATTTATCTCTCCACCACTCAACTCAAGAGCCAAAT 1981
QY 3522 ACTGATCGCAACTTTATGAAAATCTTCCAGTTCTACAAAATCTTACTGAGTTTGAAGAA 3581
Db 1982 ACGGATCGGCACTTTATGAAAATCTTCCAGTTCTACAGAAAATCTTACGAGTTTGAAGAA 2041
QY 3582 GCAATTTGGAGTAAATTTTACTCATGTTGACCTTCTGGCAAGGGCATTTACATTTGAGAACT 3641
Db 2042 GCAATTTGGAGTAAATTTTACTCATGTTGACCTTCTGGCAAGGGCGTTTACATGAGAAC 2101
QY 3642 GTGGGATTTAAACCATCTGACCCCTAGGCCCAATACAGAGAAATGGAATTCCTAGTGTGCTCC 3701
Db 2102 GTGGGCTTTAAACCATCTGACCCCTAGGCCCAATACAGAGGATGGAATTTCTGGGCGACTCC 2161
QY 3702 ATAAATGCAACTGCTGATGACAGAGTACTTTATTCATTTCTCCAGATCATCATGAGAA 3761
Db 2162 ATAAATGCACTGCTGATGACAGAGTACTTTATTCATTTCTCCAGATCATCATGAGAA 2221
QY 3762 CACTTAACTTTCTGCGAGGCTCTTTGTTGAATTAATAGAACTCAGGCCAAGGTAGCGGAG 3821
Db 2222 CACTTGAAGCTTTTGAAGTCTTTTGAAGTCTTTTGAAGTCTTTTGAAGTCTTTTGAAGTCTTT 2281
QY 3822 GAGCTGGGATGACGAGGATGACCCCAATCCAGCTTTCAGCAGTGTGTTGCTTGAACATTT 3881
Db 2282 GAACTGGGATGACGAGGATGACCCCAATCCAGCTTTCAGCAGTGTGTTGCTTGAACATTT 2341
QY 3882 CGCACCAAGACTTGGCGGACCTTTTGGAAATCATTTTATGAGCGCTGTACACTGATAG 3941
Db 2342 AGAACCAAGACTTGGCGGACCTTTTGGAAATCATTTTATGAGCGCTGTACACTGCAAG 2401
QY 3942 GATTTTGGATATGTTTCACTATGATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4001
Db 2402 GACCTGGAATATGTTCCACACTTTTCAATGAACTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2461
QY 4002 ATTTTGAATCAGGATTTGGAATGACCCCAATCCAGCTTTCAGCAGTGTGTTGCTTGAACATTT 4061
Db 2462 ATTCTGATCAGGATTTGGAACTGACCCCAAGTTCGAGCTGACGAGTGTGTTGCTTGAACATTT 2521
QY 4062 AGGACAGAAAGGAAAGAGCCAGCAATTTCTCTGTATCAAGACTCTGACAGACAGTGGGCCCA 4121
Db 2522 AGGACAGAAAGGAAAGAGCTGACACTCCCTTTATACAGACTCTGACAGACAGTGGGCCCA 2581

Qy	4122	TCCCATGCCCGAACCTACACTGTGGCTGTTTATTTCAAGGGAGAGAAAGAAATAGGCTGTGGG	4181
Db	2582	TCCCATGCTAGAACCTACACTGTGGCTGTTTATTTCAAGGGAGAGAAAGAAATAGGCTGTGGG	2641
Qy	4182	AAAGGACCAAGTATTTCAGCAAGCGGAAATGGGAGCAGCAATGGATGCGCTTGAAAAATAT	4241
Db	2642	AAAGGACCAAGCATTCAGCAGCGCGAGATGGGAGCAGCAATGGATGCGCTTGGAAGAAATAT	2701
Qy	4242	AATTTTCCCCAGATGGCCCATCAGAACGGGTTTCATCGAACCGGAAGTACAGACAAGAGTTA	4301
Db	2702	AATTTTCCCCAGATGGCCCATCAGAACGGGTTTCATCGAACCGGAAGTACAGACAAGAGTTA	2761
Qy	4302	AAAGAAATGAGGTGGGAAAGAGAGAGCATCAAGAGAGAGAGCCAGATGAGACTGAAAGACATC	4361
Db	2762	AAAGAAATGAGGTGGGAAAGAGAGAGCATCAGAGAGAGAGCCAGAGGAGGCTGAAGACATC	2821
Qy	4362	AGAAATAAAGAGGGGCATGCAAGTGTGGAGTATTTACTTGTCTCAGTAACTGTGACTGTT	4421
Db	2822	AGAAATAAAGAGGGGCAGG--AGTGGCGGAGCATTTGCCCTTTCTAAGAACTGTGACTGTG	2879
Qy	4422	GTCTATTGAGACCTAGCTAGTTTCTCTGCAGACAATGAACGAAGTGTGCTCATTGAAAT	4481
Db	2880	GCCCATCGAGACCTAGCTGGTTTTCCTTTAGACAATGAATGAAGTGTGCCCATTTGAAAT	2939
Qy	4482	AAAATACAGA	4491
Db	2940	AAACTCAAA	2949

Search completed: December 26, 2005, 22:22:23
Job time : 23229 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 00:37:50 ; Search time 265 Seconds
(without alignments)
3658.098 Million cell updates/sec

Title: US-10-774-974-2
Perfect score: 7500
Sequence: 1 MMQNTCHRMSTFPGRCPR.....MRWERHQREPDTEDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot.sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	NC	Score	Query	Match	Length	DB	ID	Description
1	7471	99.6	1374	1	RNC_HUMAN			Q9nr4 homo sapien
2	7160	95.5	1373	2	Q5HZJ0_MOUSE			Q5hzj0 mus musculu
3	6380	85.1	1231	2	Q6PAJ6_MOUSE			Q6paj6 mus musculu
4	6344.5	84.6	1336	2	Q5ZIR3_CHICK			Q5zir3 gallus gall
5	5217	69.6	1021	2	Q80Z69_MOUSE			Q80z69 mus musculu
6	4982	66.4	942	2	Q6P88_MOUSE			Q6p88 mus musculu
7	4121	54.9	780	2	Q7TM18_MOUSE			Q7tm18 mus musculu
8	3767	50.2	1041	2	Q4SKW1_TETNG			Q4skw1 tetraodon n
9	3696	49.3	705	2	Q80U05_MOUSE			Q80u05 mus musculu
10	2963	39.5	1327	2	Q9XYN5_DROME			Q9xyn5 drosophila
11	2881.5	38.4	979	2	Q5TOK8_ANOGA			Q5tok8 anopheles g
12	2825	37.7	541	2	Q8CJ74_MOUSE			Q8cj74 mus musculu
13	2811.5	37.5	1071	2	Q960Y4_DROME			Q960y4 drosophila
14	2683.5	35.8	836	2	Q7PNE9_MOUSE			Q7pne9 anopheles g
15	2298	30.6	432	2	Q59FF9_HUMAN			Q59ff9 homo sapien
16	1480	19.7	288	2	Q67B98_RAT			Q67b98 rattus norv
17	1473.5	19.6	1061	2	Q61XX9_CAEBR			Q61xx9 caenorhabdi
18	1448	19.3	1086	1	RNC_CAEEL			Q01326 caenorhabdi
19	1154.5	15.4	290	2	Q4SKW3_TETNG			Q4skw3 tetraodon n
20	1138.5	15.2	412	2	Q5G9M4_XENLA			Q5g9m4 xenopus lae
21	711	9.5	134	2	Q6CTG2_MOUSE			Q6ctg2 mus musculu
22	344.5	4.6	230	2	Q82ZG1_ENTFA			Q82zg1 enterococcu
23	340.5	4.5	1150	2	Q8INM6_DROME			Q8inm6 drosophila
24	340	4.5	846	2	Q8MRP6_DROME			Q8mrp6 drosophila
25	340	4.5	1109	2	Q9VAY4_DROME			Q9vay4 drosophila
26	339	4.5	1389	2	Q55FS1_DICDI			Q55fs1 dictyosteli
27	331	4.4	1838	2	Q6BCJ9_TETTH			Q6bcj9 tetrahymena
28	329.5	4.4	2869	2	Q60PD2_CAEBR			Q60pd2 caenorhabdi
29	328	4.4	1022	2	Q4SPX9_TETNG			Q4spx9 tetraodon n
30	325	4.3	551	2	Q8BX18_MOUSE			Q8bx18 mus musculu
31	321	4.3	249	2	Q65UQ5_BACLD			Q65jq5 bacillus li

RESULT 1									
ID	RNC_HUMAN	STANDARD;	PRT;	1374	AA.				
AC	Q9NR44; Q7Z5V2; Q86YH0; Q9NW73; Q9Y2V9; Q9Y4Y0;								Q8bx86 mus musculu
DT	28-FEB-2003 (Rel. 41, Created)								Q6nvf9 mus musculu
DT	28-FEB-2003 (Rel. 41, Last sequence update)								Q7s9h3 neurospora
DT	13-SEP-2005 (Rel. 48, Last annotation update)								Q5nv88 pongo pygma
DE	Ribonuclease III (EC 3.1.26.3) (RNase III) (Drosha) (p241).								Q53es1 homo sapien
GN	Name=RNASE3L; Synonyms=RN3;								Q16630 homo sapien
OS	Homo sapiens (Human)								P51833 bacillus su
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								Q4sq65 tetraodon n
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;								Q6ddw4 xenopus lae
OC	Homo.								Q5z134 gallus gall
OX	NCBI_TaxID=9606;								Q7xq14 oryza sativ
RN	[1]								Q819v8 bacillus ce
RP	NUCLEOTIDE SEQUENCE (ISOFORM 1).								Q8y691 listeria mo
RP	MEDLINE=20538440; PubMed=10948199; DOI=10.1074/jbc.M005494200;								Q81w18 bacillus an
RA	Wu H., Xu H., Miraglia L.J., Crooke S.T.;								
RT	"Human RNase III is a 160-kDa protein involved in preribosomal RNA								
RT	processing.";								
RL	J. Biol. Chem. 275:36957-36965(2000).								
RN	[2]								
RP	NUCLEOTIDE SEQUENCE OF 166-613 (ISOFORM 2).								
RC	TISSUE=Colon;								
RX	MEDLINE=20431278; PubMed=10976766; DOI=10.1023/A:100717623283;								
RA	Gunther M., Laithier M., Brisson O.;								
RT	"A set of proteins interacting with transcription factor Sp1								
RT	identified in a two-hybrid screening.";								
RL	Mol. Cell. Biochem. 210:131-142(2000).								
RN	[3]								
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 603-1374.								
RC	TISSUE=Embryo;								
RX	PubMed=14702039; DOI=10.1038/ng1285;								
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,								
RA	Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Makita H.,								
RA	Sekine M., Ohyasui M., Nishi T., Shibahara T., Tanaka T., Ishii S.,								
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,								
RA	Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,								
RA	Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,								
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,								
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,								
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,								
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,								
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotta T.,								
RA	Kusano J., Togehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,								
RA	Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,								
RA	Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,								
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,								
RA	Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,								
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,								
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,								
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,								
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,								
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,								

RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Saeaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3), AND NUCLEOTIDE
RP SEQUENCE [LARGE SCALE MRNA] OF 553-1374 (ISOFORM 1).
RC TISSUE=Cervix, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalak U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 706-1374.
RC TISSUE=Aorta;
RA Wei Y.J., Ding J.F., Xiong H., Zhou Y., Liew C.C.;
RA Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP FUNCTION
RX PubMed=14508493; DOI=10.1038/nature01957;
RA Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P.,
RA Raadmark O., Kim S., Kim V.N.;
RT "The nuclear RNase III Drosha initiates microRNA processing.";
RL Nature 425:415-419(2003).
CC -I- FUNCTION: Executes the initial step of microRNA (miRNA) processing
CC in the nucleus, that is cleavage of pri-miRNA to release pre-
CC miRNA. Involved in pre-rRNA processing. Cleaves double-strand RNA
CC and does not cleave single-strand RNA.
CC -I- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonooester.
CC -I- SUBUNIT: Interacts with Spl.
CC -I- INTERACTION: Q8WY05:DGCR8; NbExp=2; IntAct=EBI-528367; EBI-528411;
CC -I- SUBCELLULAR LOCATION: Nuclear. A fraction is translocated to the
CC nucleolus during the S phase of the cell cycle.
CC -I- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9NRR4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NRR4-2; Sequence=VSP_005777;
CC Note=No experimental confirmation available;
CC Name=3;
CC IsoId=Q9NRR4-3; Sequence=VSP_012450, VSP_012451, VSP_012452,
CC VSP_012453;
CC Note=No experimental confirmation available;
CC TISSUE SPECIFICITY: Ubiquitous.
CC -I- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
CC -I- SIMILARITY: Contains 2 RNase III domains.
CC -I- CAUTION: Ref.3 sequence differs from that shown due to a

CC frameshift in position 775.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF189011; AAF80558.1; -; mRNA.
CC EMBL; AJ242976; CAB45133.1; -; mRNA.
CC EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC EMBL; BC041162; AAH41162.1; -; mRNA.
CC EMBL; BC054003; AAH54003.1; -; mRNA.
CC EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -; mRNA.
CC DR EMBL; AK001121; BAA91511.1; ALT INIT; mRNA.
CC DR EMBL; BC041162; AAH41162.1; -; mRNA.
CC DR EMBL; BC054003; AAH54003.1; -; mRNA.
CC DR EMBL; AF116910; AAD29637.1; ALT_FRAME; mRNA.
CC HSSP; O67082; 1JF2.
CC -----
CC DR EMBL; AF189011; AAF80558.1; -; mRNA.
CC DR EMBL; AJ242976; CAB45133.1; -

Db	1	MQGNTCHRMVTHPGRCGRGGHGAAPSAPAFRPNQLRLHLHFQQQPAQYQYEPDPSPSS	60
Qy	62	TFSNSPAPNFLPRPDEFVPPPPMPPPSAQGLPFCPTIRPPFPNQHMRHPVPVPCFPMP	121
Db	61	SYSNSQAPSPWPRPDEFVPPPPAAQGLPFCPVPRPPYPNHMRHPVPVPCFPMP	120
Qy	122	PPMPCPNNPVPGAGPPQCGTTPPFMMPPSPMHPPPPPMPPQVNYQYPCYSHNTPPPS	181
Db	121	PPMPCPNNPASGAPPGCGTTPPFMMVPPSPMHPPPPPMPPQVNYQYPCGS-HSPFPFG	179
Qy	182	FNSFQNNPSPFLPSANNSSPHFRHLPPYPLKAPSPERRSPERLKHVDHHRHDSHGSG	241
Db	180	FNSYQNNSSFPFSPANSSSPHFHFLPPYSLPKAQNERRSPERLKHVDHHRHDSHGSG	239
Qy	242	ERHSLDRRRGRSPDRRRQDSRYRSYDRGRT-PSHRSYERSRERERHRHRDNRRS	300
Db	240	ERHSLERRRGRSPERRRPSRYSYDRGRTPPPHRSYERSRERDRHRHRREARS	299
Qy	301	PSLERSYKKEYKRGSRSGYLSVVEPAGCTPELPGELIKNTDQWAPLEELVNHRSRSREK	360
Db	300	PSLERSYKKEYKRGSRGVALVPAPEPAGCTPELPGEMIKTTESWAPPENNVNHRSPSREK	359
Qy	361	KRAEWEEEKRWSNQSGKQKNYTSIKEKEPRETMPDKNEESEEELKPEVIRCTHSEN	420
Db	360	KRAEWEEEKRWSQSGKQKNYTSIKEKEASEVPEKTEEBEEBELKPEVIRCTHSSS	419
Qy	421	YYSDDPMQVGDSTVWGTSLRDLYDXFEELSGRQEKAKAAPMPBPPTKLDDELESS	480
Db	420	YYSDDPMQVGDSTVWGTSLRDLYDXFEELGNRQEKAKAAPMPBPPTKLDDELESS	479
Qy	481	SESECEDESDSCSSSDSESVFOVIAEIKRKAHPDRLHDELWYNDPGQWMDGPLCKCSA	540
Db	480	SESECTDDSDSCSSSDSESVFVDIAEIKRKAHPDRLHDELWYNDPGQWMDGFLCKCSA	539
Qy	541	KARTGTIRHSIYQGEBAIKPCRPMNNAGRLFHYRIIVSPPTWFLTDTPPIVSYDDHEYI	600
Db	540	KARTGTIRHSIYQGEBAIKPCRPMNNAGRLFHYRIIVSPPTWFLTDTPPIVSYDDHEYI	599
Qy	601	FEGSMFAHALPTNIPLCKVIRFNIDYTHPIEBMMPENFCVKGLELFSLFRDILELY	660
Db	600	FEGSMFAHALPTNIPLCKVIRFNIDYTHPIEBMMPENFCVKGLELFSLFRDILELY	659
Qy	661	DWNLKGPLFEDSPCCPRFPMFPRFVRLPDGGKEVLSMHQIILLYLRCSSKALVPBEEIA	720
Db	660	DWNLKGPLFEDSPCCPRFPMFPRFVRLPDGGKEVLSMHQIILLYLRCSSKALVPBEEIA	719
Qy	721	NMLQWEELEWQYAECKGMIVTHPGTKPSSVRIDQLDREQFNPDVITPPIIVHFGIRPA	780
Db	720	NMLQWEELEWQYAECKGMIVTHPGTKPSSVRIDQLDREQFNPDVITPPIIVHFGIRPA	779
Qy	781	QLSVAGDPQYOKLWKSYYVKLRLHLANSPKVKQTDQKLAOREALOKIROKNTMRREVTY	840
Db	780	QLSVAGDPQYOKLWKSYYVKLRLHLANSPKVKQTDQKLAOREALOKIROKNTMRREVTY	839
Qy	841	ELSSQGFWKGTGRSDVCOHMMPLVLTTHIRYHOCLMHLDKLIGYTFQDRCLLQLAMTHP	900
Db	840	ELSSQGFWKGTGRSDVCOHMMPLVLTTHIRYHOCLMHLDKLIGYTFQDRCLLQLAMTHP	899
Qy	901	SHHLNFGMNPDPHARNSLNCGIRPKYKGDRCVKVHMHRKKGINTLINIMRSLGQDDTPPS	960
Db	900	SHHLNFGMNPDPHARNSLNCGIRPKYKGDRCVKVHMHRKKGINTLINIMRSLGQDDTPPS	959
Qy	961	RINHNRELEFLGDAAVEFLTSVHLXYLPPSLEEGGLATYRTAIVONOHLAMLAKKLELDP	1020
Db	960	RINHNRELEFLGDAAVEFLTSVHLXYLPPSLEEGGLATYRTAIVONOHLAMLAKKLELDR	1019
Qy	1021	FWLYAHGPDLCRESDLRHAMANCPEALIGAVYLSGSLEEAQKOLFGRLLFNDPDLREYWLN	1080
Db	1020	FWLYAHGPDLCRESDLRHAMANCPEALIGAVYLSGSLEEAQKOLFGRLLFNDPDLREYWLN	1079
Qy	1081	YPLHPLQLOEBNPTDQLIETSPVLQKLTPEBEATGVIFTHVRLLARAFPTLRTVGFNHLTL	1140
Db	1080	YPLHPLQLOEBNPTDQLIETSPVLQKLTPEBEATGVIFTHVRLLARAFPTLRTVGFNHLTL	1139

QY	1141	GHNORMEFLGDSIMQLVATEYLFTHPPDDHHEGLHTLLRSSLVNRTQAKVAEELGMOEYA	1200
DB	1140	GHNORMEFLGDSIMQLVATEYLFTHPPDDHHEGLHTLLRSSLVNRTQAKVAEELGMOEYA	1199
QY	1201	ITWNDKTRPVGLETKTLADLLESFIAALYTDKOLEVYHTFMNVCFFPRLKEFILNQDWD	1260
DB	1200	ITWNDKTRPVALTKTLADLLESFIAALYTDKOLEVYHTFMNVCFFPRLKEFILNQDWD	1259
QY	1261	PKSLOQCCCLTLTEGKEPDIPLYKTLQTVGSPHARTYTVAVYFKGRIICGKGPSIQQA	1320
DB	1260	PKSLOQCCCLTLTEGKEPDIPLYKTLQTVGSPHARTYTVAVYFKGRIICGKGPSIQQA	1319
QY	1321	EMGAAMDALEKYNFPQWAKHOKRIGIKYROELKEMWEREHOERPEDETEDIKK	1374
DB	1320	EMGAAMDALEKYNFPQWAKHOKRIFERKYROELKEMWEREHOERPEEABDIKK	1373

RESULT 3

Q6PAJ6_MOUSE

ID Q6PAJ6_MOUSE PRELIMINARY; PRT; 1231 AA.

AC Q6PAJ6;

DT 05-JUL-2004 (T-EMBLrel. 27, Created)

DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)

DE Nrasen protein (Fragment).

GN Name:Etoh12; Synonyms=Rnasen;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidae; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

ON [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Heish F.,

RA Diatchenko L., Marusina K., Farmer A.A., Casavant T.L., Scheetz T.E.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,

RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

FAhey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RL [2]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=C57BL/6; TISSUE=Brain;

RC SUBMITTED (OCT-2003) to the EMBL/GenBank/DBJ databases.

RG NIH MGC Project;

RL EMBL; BC060265; AAHG0265.1; -; mRNA.

DR EMBL; ENSMUSG0000022191; Mus musculus.

DR MGI; MGI:1261425; Etoh12.

DR MGI; MGI:1261425; Rnasen.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0003725; F:double-stranded RNA binding; IEA.

DR GO; GO:0004525; F:ribonuclease III activity; IEA.

DR GO; GO:0006336; P:RNA processing; IEA.

DR InterPro; IPR001159; D8_RNA_BD.

DR InterPro; IPR000999; RNase_III.

DR Pfam; PF00035; dsrm; 1.

DR Pfam; PF00636; Ribonuclease 3; 2.

DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS00142; RNASE_3_2; 2.
FT NON TER 1 1_2;
SQ SEQUENCE 1231 AA; 143573 MW; 3679FA9D87800151 CRC64;

Query Match 85.1%; Score 6380; DB 2; Length 1231;
Best Local Similarity 96.1%; Pred. No. 2.9e-303;
Matches 1184; Conservative 22; Mismatches 24; Indels 2; Gaps 2;

QY 144 FMVPPSPMPHPPPPVPMQVQVYPCGYSHHPPPPSPNSFPQNPSSFLPSANNSSPH 203
DB 1 FMVPPSPMPHPPPPVPMQVQVYPCGYSHHPPPPSPNSFPQNPSSFLPSANNSSPH 59

QY 204 FRLLPPVPLKAPSERSPERLKHDDHRRDHSHGGERHRSIDRRERGRSPDRRQDS 263
DB FRLLPPVPLKAPSERSPERLKHDDHRRDHSHGGERHRSIDRRERGRSPDRRQDS 119

QY 264 RYRSYDGRGT-PSRHRSYRSRERERHRRDNRSPSLERSYKKEYKRSGRSYGLSV 322
DB RYRSYDGRGT-PSRHRSYRSRERERHRRDNRSPSLERSYKKEYKRSGRSYGLSV 179

QY 323 VPFPAGCTPELPGHEI IKNTDSWAPPLEIVNHRSPSRKKKARWEEKDRDNQSSGDK 382
DB VPFPAGCTPELPGHEI IKNTDSWAPPLEIVNHRSPSRKKKARWEEKDRDNQSSGDK 239

QY 383 NYTSIKEKEEELSGRQEKAKAARPPWPKTKLDLEDLESSESECESDSDSCSSSDSEVF 502
DB NYTSIKEKEEELSGRQEKAKAARPPWPKTKLDLEDLESSESECESDSDSCSSSDSEVF 359

QY 503 DVIAEIKRKAHPDRDLHDLWYNDPGQMDGPKCKCSAKARRTGIRHSIYPGBEAKPCR 562
DB DVIAEIKRKAHPDRDLHDLWYNDPGQMDGPKCKCSAKARRTGIRHSIYPGBEAKPCR 419

QY 563 PMTNAGELPHYRTVSPPTNFLTDRPTVLEYDDHEVIFEGFSEFAHAPLTNIPCKVIR 622
DB PMTNAGELPHYRTVSPPTNFLTDRPTVLEYDDHEVIFEGFSEFAHAPLTNIPCKVIR 479

QY 623 FNIDYTHFTHEEMPNFCKVGLFSLFLFRDILEYDNLKGLPFDSPPCPFHFM 682
DB FNIDYTHFTHEEMPNFCKVGLFSLFLFRDILEYDNLKGLPFDSPPCPFHFM 539

QY 683 PRFVRLPDGKGVLSMHQILLYLLRCSKALVPEEETANMLQWEELEWQYAECKGMIV 742
DB PRFVRLPDGKGVLSMHQILLYLLRCSKALVPEEETANMLQWEELEWQYAECKGMIV 599

QY 743 TNFGTKPSSVRIDQLDREQNPDPVITPIIVHGIAPQALSYAGDPQYQKLWKSYYKLRLH 802
DB TNFGTKPSSVRIDQLDREQNPDPVITPIIVHGIAPQALSYAGDPQYQKLWKSYYKLRLH 659

QY 803 LLANSKPVKOTDKQKLAOREALQKIRQKNTMRREVTVLSQGFWTKGIRSDVCOHAMM 862
DB LLANSKPVKOTDKQKLAOREALQKIRQKNTMRREVTVLSQGFWTKGIRSDVCOHAMM 719

QY 863 LPVLTHIRYHQCIMHLDKLGITYTQDRCLQLLAWTHPSHLLNFGMNPDPHARNLSNCGI 922
DB LPVLTHIRYHQCIMHLDKLGITYTQDRCLQLLAWTHPSHLLNFGMNPDPHARNLSNCGI 779

QY 923 RQPKYGRKVVHMMRKKGINTLINIMSRIGQDDPTPSRINNHRELEFLGDVVEFLTSV 982
DB RQPKYGRKVVHMMRKKGINTLINIMSRIGQDDPTPSRINNHRELEFLGDVVEFLTSV 839

QY 983 HLYVLPSEEGGLATYRTAIVQNOHLAMAKKLELDPMLYAHGPDLCRESDLRHMAN 1042
DB HLYVLPSEEGGLATYRTAIVQNOHLAMAKKLELDRFMLYAHGPDLCRESDLRHMAN 899

QY 1043 CFEALIGAVYLEGSLEAKQLFGRLLFNDPDLREVWMLNYPHLPLQLQEPNTDQLIETSP 1102

DB 900 CFEALIGAVYLEGSLEAKQLFGRLLFNDPDLREVWMLNYPHLPLQLQEPNTDQLIETSP 959
QY 1103 VLQKLTFFEEAIGVIFTHVRELLARAFTRTVGNHNLTLGHNQRMFLGDSIMQLVATEYL 1162
DB VLQKLTFFEEAIGVIFTHVRELLARAFTRTVGNHNLTLGHNQRMFLGDSIMQLVATEYL 1019
QY 1163 FIHPDPHHEGLTLRLSSLVNNTQAKVABELGMOEYAITNDKTRPVGLRTKTLADLLE 1222
DB FIHPDPHHEGLTLRLSSLVNNTQAKVABELGMOEYAITNDKTRPVGLRTKTLADLLE 1079
QY 1223 SFTAALYTDKLEVYHTFMNVCFPPRLKBFILNQDNDPKSQLOQCCLTLRTEGKEPDIP 1282
DB SFTAALYTDKLEVYHTFMNVCFPPRLKBFILNQDNDPKSQLOQCCLTLRTEGKEPDIP 1139
QY 1283 LYKTLQTVGSHARTTVAVYFKGERIGCGKGPSIQQAENGAAAMDALEKYNFPQMAHQKR 1342
DB LYKTLQTVGSHARTTVAVYFKGERIGCGKGPSIQQAENGAAAMDALEKYNFPQMAHQKR 1199
QY 1343 FIOKRYQELKEMWREHOREPDETDIKK 1374
DB FIOKRYQELKEMWREHOREPDETDIKK 1231

RESULT 4
Q5ZIR3_CHICK PRELIMINARY; PRT; 1336 AA.
AC Q5ZIR3;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_24a21;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatki A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT gene function analysis."
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ270721; CAG32380.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS00142; RNASE_3_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 1336 AA; 154836 MW; 50D9A9AC53DE0CF CRC64;

Query Match 84.6%; Score 6344.5; DB 2; Length 1336;
Best Local Similarity 85.3%; Pred. No. 1.7e-301;
Matches 1171; Conservative 76; Mismatches 81; Indels 45; Gaps 5;

QY 10 MSFHQGRGPRGGHGAAPSAPSPQNLRLHHPQPPVQYQYPPSPATTFSPNSPAP 69
DB 1 MSPSHRGCGPRGGGPGARTSAQTYRPNLRLHHPQPPVQYQYPPSPATTFSPNSPAP 60

SQ	SEQUENCE	942 AA; 109445 MW; C3629E4676F6C8B CRC64;
	Query Match	66.4%; Score 4982; DB 2; Length 942;
	Best Local Similarity	98.9%; Pred. No. 4.1e-235;
	Matches 932; Conservative	5; Mismatches 5; Indels 0; Gaps 0;
QY	433	STVVGTSRLRLDYKFEELGSRQEKAKAARPPWPPKTKLDRDLSSSESECEDESDT 492
DB	1	STVVGTSRLRLDYKFEELGSRQEKAKAARPPWPPKTKLDRDLSSSESECEDESDT 60
QY	493	CSSSSDSEFVDVTAETIKRKAHPDRLDHLDLWYNDPGQNDGPKCSAKARTGIRHSY 552
DB	61	CSSSSDSEFVDVTAETIKRKAHPDRLDHLDLWYNDPGQNDGPKCSAKARTGIRHSY 120
QY	553	PGEEAIKPCRPMTNNAGRLPHYRIITVSPPTNFITDRPTVIEYDDHVEYIEGSGMFAHPL 612
DB	121	PGEEAIKPCRPMTNNAGRLPHYRIITVSPPTNFITDRPTVIEYDDHVEYIEGSGMFAHPL 180
QY	613	TNIPLCVKVIRFNIDYTHIFTEEMMPENFCVKGLELSFLFRDILELYDWNKGPFLFDS 672
DB	181	TNIPLCVKVIRFNIDYTHIFTEEMMPENFCVKGLELSFLFRDILELYDWNKGPFLFDS 240
QY	673	PPCCPRFHMFRVFLPDPGGKEVLSMHOILLYLLRCSKALVPEEIIANMLQWEELEWOK 732
DB	241	PPCCPRFHMFRVFLPDPGGKEVLSMHOILLYLLRCSKALVPEEIIANMLQWEELEWOK 300
QY	733	YABECKGMIVNPGTSPSSVRIDQDREQNPDPVITFPIIVHFGIRPAOLSVAGDPQYOK 792
DB	301	YABECKGMIVNPGTSPSSVRIDQDREQNPDPVITFPIIVHFGIRPAOLSVAGDPQYOK 360
QY	793	LWKSYYKLRLHLANSKPVQTDQKLAQREELQKIRQKNTMRREVTVLSSGFWKTGI 852
DB	361	LWKSYYKLRLHLANSKPVQTDQKLAQREELQKIRQKNTMRREVTVLSSGFWKTGI 420
QY	853	RSDVCQHAMWLPVLTTHIRYHQCMLDKLIGYTFQDRCLLQALWTHPSHLNFGMNPDI 912
DB	421	RSDVCQHAMWLPVLTTHIRYHQCMLDKLIGYTFQDRCLLQALWTHPSHLNFGMNPDI 480
QY	913	ARNLSNCGIRQPKYGRKVVHMHMKKGINTLIMSLRGDDPTPSRINNNERLEFLG 972
DB	481	ARNLSNCGIRQPKYGRKVVHMHMKKGINTLIMSLRGDDPTPSRINNNERLEFLG 540
QY	973	DAVVEFLTSVHLVYLPSPLEEGGLATRTAIVQNHLMALAKKLDPPFMYAHGPDLCR 1032
DB	541	DAVVEFLTSVHLVYLPSPLEEGGLATRTAIVQNHLMALAKKLDPPFMYAHGPDLCR 600
QY	1033	ESDLRHAMANCFEALIGAVYLEGSLEBEAKQLFGRLLFNDPDLREVWLNYPHLPLQEPN 1092
DB	601	ESDLRHAMANCFEALIGAVYLEGSLEBEAKQLFGRLLFNDPDLREVWLNYPHLPLQEPN 660
QY	1093	TDRLQIETSPVLQKLFTEFEBAIGVIFTHVLLARAFITRTVGNHLLTGHQRMFLGDS 1152
DB	661	TDRLQIETSPVLQKLFTEFEBAIGVIFTHVLLARAFITRTVGNHLLTGHQRMFLGDS 720
QY	1153	IMOLVATEYLFIFHPDHHGHTLLRSSLVNNRTOAKVABELGQVEYAINDTKTKPVGL 1212
DB	721	IMOLVATEYLFIFHPDHHGHTLLRSSLVNNRTOAKVABELGQVEYAINDTKTKPVGL 780
QY	1213	RTKTLADLLESFIAALYTDKDLVYVHTFMNVCFPPRLKEFILNQDWNDPKSQLQCCCLTL 1272
DB	781	RTKTLADLLESFIAALYTDKDLVYVHTFMNVCFPPRLKEFILNQDWNDPKSQLQCCCLTL 840
QY	1273	RTSGKGPDIPLYKTQTVGSHARTYTVAVYFGERIGCGKGPSIOQAENGAMDALEKY 1332
DB	841	RTSGKGPDIPLYKTQTVGSHARTYTVAVYFGERIGCGKGPSIOQAENGAMDALEKY 900
QY	1333	NFPQMAHQKRFGRKRYQELKEMRWREHOREPDETEDETKK 1374
DB	901	NFPQMAHQKRFGRKRYQELKEMRWREHOREPDETEDETKK 942

RESULT 7
.Q7TWI8_MOUSE

ID	Q7TWI8_MOUSE PRELIMINARY;	PRT; 780 AA.
AC	Q7TWI8;	
DT	01-OCT-2003 (trEMBLrel. 25, Created)	
DT	01-OCT-2003 (trEMBLrel. 25, Last sequence update)	
DT	01-MAR-2004 (trEMBLrel. 26, Last annotation update)	
DE	Rnaasen protein (Fragment).	
GN	Name=Etchi2; Synonyms=Rnaasen;	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;	
OC	Muroidea; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RC	STRAIN=C57BL/6; TISSUE=Mouse;	
RX	MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;	
RA	Strausberg R.L., Feilgold E.A., Grouse L.H., Derge J.G., Schuler G.D.,	
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,	
RA	Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human	
RT	and mouse cDNA sequences.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
RN	[2]	
RP	NUCLEOTIDE SEQUENCE.	
RC	STRAIN=C57BL/6; TISSUE=Mouse;	
RG	NIH MGC Project;	
RL	Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; BC055696; AAH55696.1; -; mRNA.	
DR	HSSP; O67082; 1JFZ.	
DR	Ensembl; ENSMUSG0000022191; Mus musculus.	
DR	MGI; MGI:1361425; Etchi2.	
DR	MGI; MGI:1361425; Rnaasen.	
DR	GO; GO:0003622; C:intracellular; IEA.	
DR	GO; GO:0003725; F:double-stranded RNA binding; IEA.	
DR	GO; GO:0004525; F:ribonuclease III activity; IEA.	
DR	GO; GO:0006396; P:RNA processing; IEA.	
DR	InterPro; IPR001159; Da_RNA_bd.	
DR	InterPro; IPR000999; RNase_III.	
DR	Pfam; PF00035; dsrm; 1.	
DR	Pfam; PF00636; Ribonuclease_3; 2.	
DR	SMART; SM00358; DSRM; 1.	
DR	SMART; SM00535; RIBOC; 2.	
DR	PROSITE; PS00137; DS_RBD; 1.	
DR	PROSITE; PS00517; RNase_3_1; 2.	
DR	PROSITE; PS01442; RNase_3_2; 2.	
FT	NON_TER	
SQ	SEQUENCE 780 AA; 91045 MW; 78D734332C75382E CRC64;	

Query Match	54.9%;	Score 4121;	DB 2;	Length 780;
Best Local Similarity	99.1%;	Pred. No. 3.8e-193;		
Matches 773;	Conservative 2;	Mismatches 5;	Indels 0;	Gaps 0;
Qy	595	DDHEYIFEGFSMAHAPLTNIPCKVIRFNIDYTHIFIEEMMPENFCVKGLELSFLFR	654	
Db	1	DDHEYIFEGFSMAHAPLTNIPCKVIRFNIDYTHIFIEEMMPENFCVKGLELSFLFR	60	
Qy	655	DILELYDWNKGPFLFDSPPCCPRFHMFRVFLPDGGKEVLSMHOILLYLRCSKALV	714	
Db	61	DILELYDWNKGPFLFDSPPCCPRFHMFRVFLPDGGKEVLSMHOILLYLRCSKALV	120	

QY 715 PEEIANNLOWERLEWQKAYAECKGMIVTNGTKPSVRIDQLDRQFNPDPVITPFIIVH 774
 DB 121 PEEIANNLOWERLEWQKAYAECKGMIVTNGTKPSVRIDQLDRQFNPDPVITPFIIVH 180
 QY 775 FGIRPAQLSVAGDPQYQKWLKSVVCLRHLLANSKPKVQKTDKQKLAQREELAQIKRQNTM 834
 DB 181 FGIRPAQLSVAGDPQYQKWLKSVVCLRHLLANSKPKVQKTDKQKLAQREELAQIKRQNTM 240
 QY 835 RREVTVELSSQGFWKGTGIRSDVCOHAMMLPVLTHIRYHQLMHLDKLIGYTFQDRCLLQ 894
 DB 241 RREVTVELSSQGFWKGTGIRSDVCOHAMMLPVLTHIRYHQLMHLDKLIGYTFQDRCLLQ 300
 QY 895 LAMTHSHHLNFCMNDPHARNLSNCGIRPKYGDRAKVVHMMRKKGINTLINIMSRLLQ 954
 DB 301 LAMTHSHHLNFCMNDPHARNLSNCGIRPKYGDRAKVVHMMRKKGINTLINIMSRLLQ 360
 QY 955 DDPTPSRINNHRELEFLGDAVVEFLTSVHLIYLPSPLEEGGLATYRTAIVQNOHLAMLA 1014
 DB 361 DDPTPSRINNHRELEFLGDAVVEFLTSVHLIYLPSPLEEGGLATYRTAIVQNOHLAMLA 420
 QY 1015 KLELDPFMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLLEBAKQLFGRLLFNDPDL 1074
 DB 421 KLELDPFMYAHGPDLCRESDLRHAMANCEALIGAVYLEGSLLEBAKQLFGRLLFNDPDL 480
 QY 1075 REVWLYPLHLOQEPNTDRQLIETSPVLQKLTETFEBAIGVITFVRLARAFITLRTVG 1134
 DB 481 REVWLYPLHLOQEPNTDRQLIETSPVLQKLTETFEBAIGVITFVRLARAFITLRTVG 540
 QY 1135 FNLHTLGHNMREFLGDSTMLQVATVFLTHPPDHHEGLHTLRSLSLVNVRTQAKVABE 1194
 DB 541 FNLHTLGHNMREFLGDSTMLQVATVFLTHPPDHHEGLHTLRSLSLVNVRTQAKVABE 600
 QY 1195 GMDQYAITNDKTRPVGLRTKTLADLLESFIAALYTDKOLEYVHTFMNVCFPPRLKEFIL 1254
 DB 601 GMDQYAITNDKTRPVGLRTKTLADLLESFIAALYTDKOLEYVHTFMNVCFPPRLKEFIL 660
 QY 1255 NQDNDPKSQLOCCLTLTREGKEPDIPLYKTQTVGPSHARTYTVAVTFKGRIGCGKG 1314
 DB 661 NQDNDPKSQLOCCLTLTREGKEPDIPLYKTQTVGPSHARTYTVAVTFKGRIGCGKG 720
 QY 1315 PSTQQAEMGAAMDALKETNPPMAHOKRFTGRKYROELKEMRERHOREPDETEDIK 1374
 DB 721 PSTQQAEMGAAMDALKETNPPMAHOKRFTGRKYROELKEMRERHOREPDEAEDIK 780
 RESULT 8
 Q4SKW1_TETNG PRELIMINARY; PRT; 1041 AA.
 AC Q4SKW1;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
 DE Chromosome undetermined SCAFL4565, whole genome shotgun sequence.
 GN ORNames=GSTENG001653001;
 OS Tetraodon nigroviridis (green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OK NCBI_Taxid=99883;
 RN NUCLEOTIDE SEQUENCE.
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Coataz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthonard V., Jubin C., Castell V., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
 RA Parra G., Lardier G., Chappelle C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo B., Nusbaum C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Winkler P., Iander E.S., Weissbach J., Roest Crolius H.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype."; Nature 431:946-957(2004).
 RL Nature 431:946-957(2004).
 RN NUCLEOTIDE SEQUENCE.
 RP Genoscope; Whitehead Institute Centre for Genome Research;
 RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 RL -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CAE01014565; CAF98721.1; -: Genomic DNA.
 SQ SEQUENCE 1041 AA; 120534 MW; 2CA7A165A7BB3CDB CRC64;
 Query Match 50.2%; Score 3767; DB 2; Length 1041;
 Best Local Similarity 67.5%; Pred. No. 1.1e-175;
 Matches 738; Conservative 80; Mismatches 135; Indels 140; Gaps 18;
 QY 10 MSFHPGCGPCRGCGHGAAPSAPSPQNLRLHLPQPPVQYQYEPSPASTTF-SNSPA 68
 DB 1 MSFHQDRAAHRFQMAFGPHMSQNTYPTNLRPQPVPVPPHYHYDQVPSSGYHSNS-- 58
 QY 69 PNFLLPRPDPVFPFP-----PMPPSAQGP-LPPCPIRPPFPNHPHPPFPVPPC 116
 DB 59 -GTMHPHSTFMQVSPGAFADSRSEPVPARQGGPHLPCLPRTPTSLYMMHPVPPPLA 117
 QY 117 FPPMPMPMPCPN---NPPVPGAPPGQGTFFFMVPPSPMHPPPPPVMPQOVNYQYPPGS 173
 DB 118 RPPMPAPIHSPMSYHPPY-----PMNSPHPPHP----- 147
 QY 174 HNFPPPPSPNS-FQNNPSSFLPSANNSSPHFPHLPYPPLPKAPSRERRSPERLKHVDDHR 232
 DB 148 ---FPPPTFNSGHVQSPGSAFDLPG-----FHHGGFYKHKDPADKRRSPERGHRDDL 198
 QY 233 HRDHS-GRGERHSL---DRRERGRSPDR--RRDGRVRSYDRCRTSRHRSYSRSRE 286
 DB 199 QGYSHSGHSDKQMEFFGDRKDRGSPDRRWPEGRHRSEYDRGTTPRH-----RSRE 254
 QY 287 RERERHRHNRDRSPSLERSYKKEYKRSQSGYGLSVVPEPAGCTPELPGELIKNTDSWAP 346
 DB 255 RSERYVHRDRSRSSQSPDRHRKRPSRSG----- 283
 QY 347 PLIEVNHRSRREKGRARWEEKDRWSDNQSS-GDKNTYSIKEKEPEETMPDK----- 399
 DB 284 -----SRERKSRWEEERERRSSAPSRRERSYVSARNRDSSEAVPERERGGFP 333
 QY 400 -----NEEEEEEELKPVWIRCTHSENYSYSDPMDQVGDSTVVG 438
 DB 334 TGRGEGEPALPEHREKE 393
 QY 439 SRLRLDYKFEBELASRQEKAKAAPPPWPPPKTKLDEDL-ESSSESECESED-STCSS 496
 DB 394 SKLRDLYQFEELGFRQARAKAAPKWEPPKTKLDEDPDESSESECESEDADGSCSS 453
 QY 497 SDESEVDVTAIEKRAKHAHDLRLHLDLWYNDPGMNDGPKCKSAKARRTGIHRSIYPGEE 556
 DB 454 SDDSDVDVTAIEKRAKHAHDLRLHLDLWYNDPGMNDGPKCKSAKARRTGIHRSIYPGEE 513
 QY 557 AIKPCPEMTNAGRLPHYRITVSPPTNFLTDRPTVLEYDDHEVIFEGFSWFAHPLTNIP 616
 DB 514 SVKQCRAMNNAGKLFPHYRITVSPPTNFLTDRPTVLEYDDHEVIFEGFSWFAHPLTNIP 573
 QY 617 LCKVIRFNIDYTHIFTEEMMPENFCVKGLFELSLFLDRDILEYDNLNKGFLPEDSPCC 676
 DB 574 LCRVIRFNIDYTHIFTEEMTPENYCVRGLELSSYLFDQILEYDNLNKGFLPDSDSPGC 633
 QY 677 PRFHFMFRVRLPGGKEVLSHQILLYLLRCSKALVPEEEIANMLWEELEWQYAE 736
 DB 634 PQFHFMFRVRLPGGKEVLSHQVLLYLLRCSKALVPEEEIANMLWEELEWQYAE 693
 QY 737 CKGMIVTNGTKPSVRIDQLDRQFNPDPVITPFIIVHFGIRPAQLSVAGDPQYQKWLK 796

694 CKGMIVTNGMKPSSVRIDQDREQFNPVDVITPPIIVHFGIRPAQLSYAGDPQYQKLMKS 753

797 YVKLRHLLANS PKVKTQDKQKLAOREALQKIQKNTMRREVTVLSSQGFWKTGIRSDV 856

754 YVKLRHLLANS PKVKTQDKQKLMOREALQKIQKNTMRREVTVLSSQGFWKSGIRSDV 813

857 CQHAWMLPVLTHIRYHQCMLHLDKLGIVTFQDRCCLQLQAMTHPSHLLNFGWNPDHARNS 916

814 CQHAWMLPVLTHIRYHQCMLHLDKLGIVTFQDRCCLQLQAMTHPSHLLNFGWNPDHARNS 873

917 LSNCGIRQPKYGDVKVHHMHRKKGINTLINIMSLRGQDDPTPSRINHNRLERLEFLGDVY 976

874 LSNCGIRQPKYGDVKVHHMVRKKGINTLINIMSLRGQDDPSRINHNRLERLEFLGDVY 933

977 EFLTSVHLVYLPSPLEEGGLATYRTAIVONQHMLAKKLELDPFMYAHGDPDLCSDDL 1036

934 EFLTSVHLVYLPSPLEEGGLATYRTAIVONQHMLAKKLELDRFMLYAHGDPDLCSDDL 993

1037 RHAWANCFEALIG 1049

994 RHAWANCFEALIG 1006

RESULT 9

Q80UUS MOUSE

ID Q80UUS_MOUSE PRELIMINARY; PRT; 705 AA.

AC Q80UUS;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Rnaasen Protein (Fragment).

GN Names:Etohi2; Synonyms:Rnaasen;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidae; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=ES Cell;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S., Krzywinski M.I., Skaleka U., Smalilus D.E.,

RA "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=129/Sv x 129/Sv-CP; TISSUE=ES Cell;

RG NTH MGC Project;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

DR ENBL; BC050057; AAH50057.1; -, mRNA.

DR HSSP; O67082; tJFZ.

DR Ensembl; ENSMUSG00000022191; Mus musculus.

DR MGI; MGI:1261425; Etohi2.

DR MGI; MGI:1261425; Rnaasen.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0003725; P:double-stranded RNA binding; IEA.

DR GO; GO:0004525; F:ribonuclease III activity; IEA.

DR GO; GO:0006396; P:RNA processing; IEA.

DR InterPro; IPR001159; Ds_RNA_bd.

DR InterPro; IPR000999; RNase_III.

DR Pfam; PF00035; dsrm; 1.

DR SMART; SM00358; Ribonuclease_3; 2.

DR SMART; SM00535; DSRM; 1.

DR PROSITE; PS0137; DS_RBD; 1.

DR PROSITE; PS00517; RNase_3_1; 2.

DR PROSITE; PS0142; RNase_3_2; 2.

FT NON_TER 1

ST SEQUENCE 705 AA; 82084 MW; 3332C69A5F1977D6 CRC64;

Query Match 49.3%; Score 3696; DB 2; Length 705;

Best Local Similarity 98.7%; Pred. No. 2e-172; Indels 0; Gaps 0;

Matches 696; Conservative 2; Mismatches 7;

QY 670 EDSPPCCPRFHPMPRVFRLPDGKKEVLSMHQILLYLLRCSKALVPEEIIANMLQWEELE 729

DB 1 EDSPPCCPRFHPMPRVFRLPDGKKEVLSMHQILLYLLRCSKALVPEEIIANMLQWEELE 60

QY 730 WQKYAECKGMIVTNPPTGTPSSVRIDQDREQFNPVDVITPPIIVHFGIRPAQLSYAGDPQ 789

DB 61 WQKYAECKGMIVTNPPTGTPSSVRIDQDREQFNPVDVITPPIIVHFGIRPAQLSYAGDPQ 120

QY 790 YOKLWKSYYKLRHLLANS PKVKTQDKQKLAOREALQKIQKNTMRREVTVLSSQGFWK 849

DB 121 YOKLWKSYYKLRHLLANS PKVKTQDKQKLAOREALQKIQKNTMRREVTVLSSQGFWK 180

QY 850 TGIKSDVDCQHAWMLPVLTHIRYHQCMLHLDKLGIVTFQDRCCLQLQAMTHPSHLLNFGMN 909

DB 181 TGIKSDVDCQHAWMLPVLTHIRYHQCMLHLDKLGIVTFQDRCCLQLQAMTHPSHLLNFGMN 240

QY 910 PHARNLSNCGIRQPKYGDVKVHHMHRKKGINTLINIMSLRGQDDPTPSRINHNRLERLE 969

DB 241 PHARNLSNCGIRQPKYGDVKVHHMHRKKGINTLINIMSLRGQDDPTPSRINHNRLERLE 300

QY 970 FLGDVAVVEFLTSVHLVYLPSPLEEGGLATYRTAIVONQHMLAKKLELDPFMYAHGPD 1029

DB 301 FLGDVAVVEFLTSVHLVYLPSPLEEGGLATYRTAIVONQHMLAKKLELDRFMLYAHGPD 360

QY 1030 LCRESDLRHAMANCFEALIGAVVLEGSLBEAKQLFGRLFNDDPDLREVMLNYPFLHPLQLQ 1089

DB 361 LCRESDLRHAMANCFEALIGAVVLEGSLBEAKQLFGRLFNDDPDLREVMLNYPFLHPLQLQ 420

QY 1090 EPNTRDQLIETSPVLOKLTPEEAIGVITFHVRLARAFITRTVGFNHLTLGHNRMEFL 1149

DB 421 EPNTRDQLIETSPVLOKLTPEEAIGVITFHVRLARAFITRTVGFNHLTLGHNRMEFL 480

QY 1150 GDSIMOLVATEYLFIHPDHPHGHGLTLRSSLVNNRTQAKVAEELGQVETINDTKTRP 1209

DB 481 GDSIMOLVATEYLFIHPDHPHGHGLTLRSSLVNNRTQAKVAEELGQVETINDTKTRP 540

QY 1210 VGLRTKTLADLLESFIAALYTDKDEYVHTFMNVCFPRPKKEFILNODMNDPKSQLOQCC 1269

DB 541 VALRTKTLADLLESFIAALYTDKDEYVHTFMNVCFPRPKKEFILNODMNDPKSQLOQCC 600

QY 1270 LTLRTGKGPDIPLYKTLQTVGSPSHARTYTVAVYFGERIGCGKGSIOQAENGAMDAL 1329

DB 601 LTLRTGKGPDIPLYKTLQTVGSPSHARTYTVAVYFGERIGCGKGSIOQAENGAMDAL 660

QY 1330 EKNYFPQMAHOKRFIERKYRQELKEMRWEREHOREPEDETIKK 1374

DB 661 EKNYFPQMAHOKRFIERKYRQELKEMRWEREHOREPEDETIKK 705

RESULT 10

Q9XYN5_DROME

ID Q9XYN5_DROME PRELIMINARY; PRT; 1327 AA.

AC Q9XYN5;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

Db 47 SSKSLDYVYPETPAPYASS-----VPSYDPYQPPAYGYEGYAVNEQAQKYGGOESHYQYQ 103
Qy 180 PSFNSQNNPSSPLPANNNSSPHRPLPPYL-----PKAPERRSPERLKHVDDHHRD 235
Db 104 PA-----SGSSFLYES-----YKYPDRYPAYSSNYRPPSERQ-----RYTS 139
Qy 236 HSHGRGERHSLDRRGRGSPDRRRQDSRYSDYDRGRTPSRHSYERSGRERERHRHR 295
Db 140 NSSQGYHYPP-----GYSSGR-----RYEQHDQ-----EHRQDSRYAEPHGHY 183
Qy 296 DNRSPSLERSYKKEYKRGSRSGYGLSVPEPAGCTPELPCEIINKTDSWAPPLEIVNHR 355
Db 184 AHRQAKGSQHY---YGSARN-----QVSDDYSPRGHERERN 219
Qy 356 PSREKKRA--RWEEEDR-----WSDNQSGDKDNYTSIKEKEPEETMPDKNEEBELK 409
Db 220 ETLEKTRAKPKVETERDRLLRQWCNFC-----EKPEDYVKGONALSEADAPV 267
Qy 410 PWMIRCTHSENYYSDDPMQVGDSTVVGTSRLRDLYDKFEELGSRQEKAKAARPPWEP 469
Db 268 ESWSRSPAELYVERTK-----SENEVRGARLQKLTFLDDELLQAKRVREKLPVYVP 323
Qy 470 KTKL-----DEDLSSSECESEDSTCSSSDSEVFDV-----IABIKRKAH 514
Db 324 PKARRRVCKHKHSEACSSSSDDSDDEDA-----FKIEDCCMEELSRKVQH 373
Qy 515 PDLHDLWYNDGQWGDPLCKCSAKARTGIRHSIYGEBAIKPCRPMTNNAAGLFHY 574
Db 374 PORVHADLWHDAGENDGPLCRCSAKSRIRGIRHGIYGETGYKLCDPNNSNAGLPHY 433
Qy 575 RITVSPPTFLTRPRTVIEDDEHYIEFGSPMAHAPLTNIPCKVIRFNIDYTIHIEB 634
Db 434 RISISPTFLTKPTIIRKHDEHEFLFEGFSLSHVRLSDPLPCKVIRFNIEYTIYEE 493
Qy 635 MPENFCVGLLEFLSLFRDILELYDNILKGPL-----FEDSPCCPRFHFMPFRVFLP 690
Db 494 KMPENTFIHLEDFIKYLFHELELVDFNLMPNLPSGNVEES---CPAFHFPPFRVRLP 550
Qy 691 DQGEVLSMHQIILYLLRSKALVPDEETANMLQWEELEWQKAECKGMVITNRTKES 750
Db 551 DNGKEVLAMVEVLYLLDNSAQIYERQQLHLNQISQSEWQNVYDFIKGLVTKPGYKC 610
Qy 751 SVRIDOLDRQENPDV-----TFPTIIVHFGIRPAQLSVAGDPQYQKLWKSYYVLR 801
Db 611 SLRVDQLDRN--NSDPEUCVDRRTGISHPAIVHFGICHPLQSLVAGPEYQKAWREVYKR 668
Qy 802 HILANSPKVQTDKQKLAQREELQKIRQKNTWRRREVTVELSSQGFWKGTIRSDVCQH 861
Db 669 HLMANMSKPSFKDKRKLKEEQRLQEMRTQGRMKRNTITVAISSEGFRTGIMCDVVQH 728
Qy 862 MLPVLTHIRYHQLMHLDKLIGYTTQDRCLLQALMTHPSHHLNFQGNPDHARNSLNC 921
Db 729 LIPVLATGLHAFKSLDLLESISIGYRFKNRYLLQALTHPSYKENYGTNPDHARNSLT 788
Qy 922 IQPKYCDRKVHMHWKKGINTLINIMSLGODDPTPSRINERLEPLGDAVVEFLTS 981
Db 789 IQPEYGDRIKXITNTKRGINTLVSIMSRFGKEHETVSNITNERLEPLGDAVVEFLS 848
Qy 982 VHLVYLPSPSEEGLATYRTAIVQNOHLAMLAKKLELDPFMLYAHGPDLCRESDLRH 1041
Db 849 IHLFFWPELEEGGLATYRAAIVQNHALLAKKLQLEEFMLYAHGSDLCHELELRH 908
Qy 1042 NCFEALIGAVYLSGEAEAKQLFGRLLF--NDPDLREVWLNYPPLHQLQBPNTDRLIET 1100
Db 909 NCFEALMGALLDGGIKVADEVFTDALFRQDEKLLSIWKNLPEHPHQEQEPGLDRSC 968
Qy 1101 SPVLQKLTETEAEAGVLFTHVRLARAFTRTVGFVHNLTIAGHORMEFLGDSITMQLV 1160
Db 969 YRVLKELTKFEDSIGIKFPHIRLLARAFTRDSIGFTHLITGSSQRLFEFLGDTVLQ 1028
Qy 1161 YLFTHFDDHHEGHLTLILRSSLVNNRTQAKVAEEIGMOEYAI--TNDKTKRPVGLRTK 1219
Db 1029 YLRHPEHHEGHLTLILRSSLVNNRTQAVVCDLGMPPKYAVYANPK---ADLKTORAD 1084

Qy 1220 LLESFIAALYTDKLEYVHTFMNVCFPPRLKEFILNQDWNPKSQLOQCCLTLRT--EGKE 1278
Db 1085 LLEAFGLALYVDKGLLYCEQFCHVCLFPLQLFINNQDWNPKSKLQCCCLTLRTMDGGE 1144
Qy 1279 PDIPYKTLQTVGPHSHARTYTVAVYFKGERICCGKGPSPSQQAEMGAMDALEKYN--FPQ 1336
Db 1145 PDIPYKVVVEASGPTNTRVYKVAVYFRSKRLATSSGSSITQQAEMNAAKQALENSRDLFPQ 1204
Qy 1337 MAHOKRFICGRYQREL---KEMRWEREHQEREPDEDEDIKK 1374
Db 1205 LDHOKRVITAKSIKKTGTGNELDNDSRQHOE-----EKIKR 1239

RESULT 11

QSTQK8 ANOQA PRELIMINARY; PRT; 979 AA.
AC QSTQK8;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE ENSANGP00000028563 (Fragment).
GN ORFNames=ENSANGG00000009133;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OX Anophelinae; Anopheles.
NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008964; BAL39656.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF000035; dsrm; 1.
DR Pfam; PF00036; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNase_3_1; 1.
DR PROSITE; PS0142; RNase_3_2; 2.
FT NON_TER
FT NON_TER
SQ SEQUENCE 979 AA; 113068 MW; 46F4AE2A0257AD92 CRC64;

Query Match 38.4%; Score 2881.5; DB 2; Length 979;
Best Local Similarity 55.3%; Pred. No. 18e-132;
Matches 546; Conservative 169; Mismatches 235; Indels 37; Gaps 12;
Qy 382 KNTYSIKKEPEETMPDKNEEBELKPKVIRCTHSENYSDPMDQVGDSTVVGTSRL 441
Db 13 RNYCETSEDIARLAEWMANDED-----RTLWVSSPAELYK-----RVSDKVVESLAR 62
Qy 442 RDIYDFEELGSRQKAKAARPPWEPPTKLTDELESSECESEDSTCSSSDSEV 501
Db 63 DALCTLFEELIKRAIRIRATQTFYPPPPRKRKMKCRHKDKC-----SSSESSEDEM 117

QY	502	-FD---VIAIEIKKKAHPDLRDLHDLWYNDPGQNDGPLCKCSAKARRTGIRHSIYPGERA	557
Db	118	EFEDCSMBELTAKIHPDLRLHVDLWYNDPGEMNDGPLCKCSARRTGIRHGKYPGEG	177
QY	558	IKPCREMTNAGRLHYRTISPTNFLTDRPTVIEYDDHEVIFEGSFMAFAPLNTIPL	617
Db	178	FKCIPNSNADKLYHRTISPTNFLTPTIIRKHDQHEFLFEGSLLAHEPIGELPT	237
QY	618	CKVIRNIDYTHIFEMBPENFCVGLBELPSLFLPRDILEYDMLKGPLFEDSPCCP	677
Db	238	CKVIRENIETIYIEQMPENFTIRELNLFDRLFRELLELVDFTVQ-PSGSGEDSSCP	296
QY	678	RHEMRFRVFLPDGCKEVLSHQIILLYLLRCSKALVPREEIANMLQWELEWQKYAEC	737
Db	297	CYHFLFRFRVDLPDNGKEVLAMSEVLRYLLDNGSLVPDPMLKEMMDMSQNEQDXYDV	356
QY	738	KGMIVTNPCKSSVRIDQLDR-----EQNPDVITFPITVHFIRGIRPAQLSYAGDPQ	789
Db	357	KGNVSNPGMKCSVRVDQLDRNVGDVPEANAIDENGLVHPVIVHFGIRPQLSYAGNEP	416
QY	790	YOKLWKSYYKLRLHLANSFKVQTDKOKLAQREALOKIRQKNTMRREVTVLSQGFWK	849
Db	417	YQKAWREYIKFRLHIANMKSPEDKRKLKAEKENRLEEMRMQGRMKNITIAVSAKAFH	476
QY	850	TGIRSDVCOHAMMLPVLTTHIRYHQCIMHLDKLIYTFQDRCLLOLANTHPSHLNFGN	909
Db	477	TGIMCDMVQHAMLIPVLTGHLFRHSANLVIRYIGYFTNRYTQLALTHPSYKENFGTN	536
QY	910	PHARNSLNCGRQPKYQDKRVHMHMKKGINTLINIMSLRGQDDTPSRINHNRL	969
Db	537	PHARNSLNCGRQPKYQDKRVHMHMKKGINTLINIMSLRGQDDTPSRINHNRL	969
QY	970	FLGDVVEFLTSVHLTYLPPSLEEGGLATYRTAIVQNHMLAKKLELDPFMYLHAGPD	1029
Db	597	FLGDVVEFLTSVHLTYLPPSLEEGGLATYRTAIVQNHMLAKKLELDPFMYLHAGPD	1029
QY	1030	LCRESDLHAMANCFALGAVYLEGSLEAKQLFGRLLFNDDP-LREVWLNYPHPLQL	1088
Db	657	LCHELELHALANCFALGAVYLEGSLEAKQLFGRLLFNDDP-LREVWLNYPHPLQL	1088
QY	1089	QEPNTDRLQIETSPVLOKLTERRAIGVIFTHVRLARAFRTLVGFNHLTLGHNMREF	1148
Db	717	QEPNTDRLQIETSPVLOKLTERRAIGVIFTHVRLARAFRTLVGFNHLTLGHNMREF	1148
QY	1149	LGDSIMQVATEYLFTHFPDHEHGLTLRSLSSVNNRTQAKVAEELGMOEYAI-TNDKTK	1207
Db	777	LGDTVLQLICSEVLYRHFPEHGHLSLRLSSVNNRTQAVVCDLGMQYAVYSNPK--	834
QY	1208	RPVGLRTKTLADLLESFIAALYTDKOLEYVHTFMNVCFPPRLKEFTLNQDNDPKSQLOQ	1267
Db	835	--ADLTKDORADLLEAFGLAYVDKGLYCEMFCVCLPRLQDFTMNQDNDPKSKLOQ	892
QY	1268	CCLTLAT-BGKEDIPLYKTQTVGSHARTYVAVYFKGERIGCGKPSIOAEMGAAM	1326
Db	893	CCLTLATMGGEPIPIVYKIECTGNTNRYSVAVYFRKRLACADGHSIOAEMNAAK	952
QY	1327	DALE--KYANFPQAHOKRFTGRKYRQ	1351
Db	953	QALENSKDLFPQLDHQRRVIAQSLKQ	979
RESULT 12			
Q8CJ74	MOUSE		
ID	Q8CJ74_MOUSE	PRELIMINARY;	PRT; 541 AA.
AC	Q8CJ74;		
DT	01-MAR-2003	(T-EMBLrel. 23, Created)	
DT	01-MAR-2003	(T-EMBLrel. 23, Last sequence update)	
DT	01-OCT-2003	(T-EMBLrel. 25, Last annotation update)	
DE	Ribonuclease III.		
GN	Name-Etchi2; Synonyms-Rn3, Rnasen;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		

RESULT 13

OC	Muroidea; Muridae; Murinae; Mus.		
OX	NCBI_TaxID-10090;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=BALB/c; TISSUE=Brain;		
RA	Min K.-L., Galarneau A., Parniak M.A.;		
RL	Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF440766; RAN63637.1; -, mRNA.		
DR	HSSP; O67082; IJFZ.		
DR	Ensembl; ENSMUSG0000022191; Mus musculus.		
DR	MGI; MGI:1261425; Etchi2.		
DR	MGI; MGI:1261425; Rnasen.		
DR	GO; GO:0005622; C:intracellular; IEA.		
DR	GO; GO:0003725; F:double-stranded RNA binding; IEA.		
DR	GO; GO:0004525; F:ribonuclease III activity; IEA.		
DR	GO; GO:0006396; P:RNA processing; IEA.		
DR	InterPro; IPR001159; Ds_RNA_bd.		
DR	InterPro; IPR000999; RNaase_III.		
DR	Pfam; PF00035; dsm; 1.		
DR	Pfam; PF00636; Ribonuclease_3; 2.		
DR	SMART; SM00358; DSRM; 1.		
DR	SMART; SM00535; RIBOG; 2.		
DR	PROSITE; PS00137; DS_RBD; 1.		
DR	PROSITE; PS00517; RNASE_3_1; 2.		
DR	PROSITE; PS0142; RNASE_3_2; 2.		
SQ	SEQUENCE 541 AA; 63007 MW; DOAC652A0CE044F1 CRC64;		
Query Match 37.7%; Score 2825; DB 2; Length 541;			
Best Local Similarity 98.3%; Pred. No. 5e-130;			
Matches 532; Conservative 3; Mismatches 6; Indels 0; Gaps 0;			
QY	834	MRREVTVLSQGFNKTGIRSDVCOHAMMLPVLTTHIRYHQCIMHLDKLIYTFQDRCLL	893
Db	1	MRREVTVLSQGFNKTGIRSDVCOHAMMLPVLTTHIRYHQCIMHLDKLIYTFQDRCLL	60
QY	894	QLAMTHPSHLNFGMNPDHARNSLNCGRQPKYQDKRVHMHMKKGINTLINIMSLRG	953
Db	61	QLAMTHPSHLNFGMNPDHARNSLNCGRQPKYQDKRVHMHMKKGINTLINIMSLRG	120
QY	954	QDDTPSRINHNRLERLEFLGDVAVEFLTSVHLTYLPPSLEEGGLATYRTAIVQNHMLA	1013
Db	121	QDDTPSRINHNRLERLEFLGDVAVEFLTSVHLTYLPPSLEEGGLATYRTAIVQNHMLA	180
QY	1014	KKLELDPFMYLHAGPDLCRESDLHAMANCFALGAVYLEGSLEAKQLFGRLLFNDDP	1073
Db	181	KKLELDPFMYLHAGPDLCRESDLHAMANCFALGAVYLEGSLEAKQLFGRLLFNDDP	240
QY	1074	LREVWLNYPHPLQLQEPNTDRLQIETSPVLOKLTERRAIGVIFTHVRLARAFRTLV	1133
Db	241	LREVWLNYPHPLQLQEPNTDRLQIETSPVLOKLTERRAIGVIFTHVRLARAFRTLV	300
QY	1134	GFNHLTLGHNMREFLGDSIMQVATEYLFTHFPDHEHGLTLRSLSSVNNRTQAKVAE	1193
Db	301	GFNHLTLGHNMREFLGDSIMQVATEYLFTHFPDHEHGLTLRSLSSVNNRTQAKVAE	360
QY	1194	LGQSYAIVTNDKTRPVGRLTKTLADLLESFIAALYTDKOLEYVHTFMNVCFPPRLKEFI	1253
Db	361	LGQSYAIVTNDKTRPVALRTKTLADLWESFIAALYTDKOLEYVHTFMNVCFPPRLKEFI	420
QY	1254	LNQDNDPKSQLOQCCLTLRTGSKBDIPLYKTQTVGSHARTYVAVYFKGERIGCGK	1313
Db	421	LNQDNDPKSQLOQCCLTLRTGSKBDIPLYKTQTVGSHARTYVAVYFKGERIGCGK	480
QY	1314	GPSIOAEMGAAMDALEKYNFPQAHOKRFTGRKYRQELKEMWEREHOREPEDETDIK	1373
Db	481	GPSIOAEMGAAMDALEKYNFPQAHOKRFTGRKYRQELKEMWEREHOREPEDETDIK	540
QY	1374	K 1374	
Db	541	K 541	

```
Q960Y4 DROME
ID Q960Y4_DROME PRELIMINARY; PRT; 1071 AA.
AC Q960Y4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LD2995p.
GN Name-drosheha; ORFNames=CG8730;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Nephrolepidae; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051775; AAK93199.1; -, mRNA.
DR HSSP; O67082; 1JFZ.
DR Ensemble; CG8730; Drosophila melanogaster.
DR FlyBase; FBgn0026722; CG8730.
DR FlyBase; FBgn0026722; drosheha.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000989; RNase_III.
DR Pfam; PF00035; dsxm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 1.
DR PROSITE; PS50142; RNASE_3_2; 2.
SQ SEQUENCE 1071 AA; 123158 MW; 25B6C2AAADA5078B CRC64;

Query Match 37.5%; Score 2811.5; DB 2; Length 1071;
Best Local Similarity 54.6%; Pred. No. 5.3e-129;
Matches 545; Conservative 169; Mismatches 220; Indels 65; Gaps 15;

QY 412 WIRCTHSENYSDPMDQVGDSTVVGTSRLRDLYDKFEELSGRQEKAKAARPPWPPKT 471
DB 14 WVRSSPAELYERTK-----SENEVRGARLQKLTFLDELLQRAKRVREKLPVYVPPR 69
QY 472 KL-----DEDLESSECEDESDSTCSSSDSEVFDV-----TAEIKRKKKAHPD 516
DB 70 KARRRVCKHKHKEACSSSSSSDDSDSDA-----FKIEQCCMBELSKVKQHPQ 119
QY 517 RLHDEIWDNDPGQWMDGPKCKSAKARTGIRHSIYPGEBAIKPCPTWNTNAGRLFYRI 576
DB 120 RVHADLWHDNDGEMNDGPKCRSAKRRIRGIRGIYPGETGYKLCDFNSNNAKGLFYRI 179
QY 577 TVGSPPTNFLTDRPTVLEYDDHEVIFEGFSMAFAPLTNIPCLKVIRFNIDYTHFTBEMW 636
DB 180 STSGPTNFLTPTTIKHDHFEFLFEGFSLSHVRLSDLPVCKVIRFNIEYTYEIEEKM 239
QY 637 PENFCVKGLFLFLFRDILEYLDWNLKGPL----FEDSPCCPRFHPFRFVRLPDG 692
DB 240 PENFTIHLDIFPKYLFHEHLELVDFNLMPNLPSGNVEES--CPAFHFFPRFVRLPDN 296
QY 693 GKEVLMSHQILLYLLRCSKALVPEEIEANMLQWEELEWQYAECKGMIVTNFGTKPSV 752
DB 297 GKEVLAMVEVRLYLLDONSQAQLVERQQLLNQISQSEWQNYVDIFKGLMVTKEGKPCSL 356
QY 753 RIDQLDREQNPDV-----ITPTLIIVHFGIRPAQLSYAGDPQYQKWLKSYVVKRLHL 803

357 RVDQLDRN--NSDLPECVDRETGISHPAIVHFGICHQPLSYAGNPEYQKAWREYVYKRL 414
804 LANSKPKVQTDKQKLAOREALQKIQKTKMREVTVELSSQGFWKTRSDVCQHAML 863
415 MANMSKSPFKDKRLEKEQRLEKQRMRTQGRMKRNTITVAISSEGFYRTGMCDVVQHAMI 474
864 PVLTHIRYHQCIMLHDLKLGITTFODRCLLQALMTHTPSHLNFGMPDPHARNSLSNCGIR 923
475 PVLTHIRYHQCIMLHDLKLGITTFODRCLLQALMTHTPSHLNFGMPDPHARNSLSNCGIR 934
924 QPKYGDGRKVVHMHMRKKGINTLINMSRLGQDDPTPSRINHNERLEFLGDVAVVEFTSVH 983
535 QPEYGDGRKIVHNTKRGINTLVSIMSRFGKEHETVSNITNERLEFLGDVAVVEFLSSIH 594
984 LYLTPSLBEGGLATYRTAIVQNHMLAKLELDPFMYAHAGPDLCHRESDLRHAWANC 1043
595 LFFMPPELEEGGLATYRAAIVQNHMLAKLELDPFMYAHAGSGLCHELELRHAWANC 654
1044 FEALIGAVYLEGSLEAKOLFGRLLF--NDPDLREVWLNYPHLPLQLOQEPNTDRLIETSP 1102
655 FEALMGALLDGGIKVADVFDDALFRQDEKLLSIWKNLPEHPLQEQEPLGDRSCIDSYR 714
1103 VLQKLTPEEEAIGVITFHVRLARAFTRTVGFNHLTLGHQNMREFLGDSIMQVATEYL 1162
715 VLKELTKFEDSIGIKFKHIRLLARAFTRDSIGTHTLTLGSKQRLFLGDTVLQLICSEYL 774
1163 FHPDHHGHHTLLRSSLVNRTQKVAEELGMQEYAT--TNDKTKRPGVLRKTLADLL 1221
775 YRHFPHEHGHLSLLRSSLVNRTQVQCDLGMKPYAVYANPK----ADLTCKORADLL 830
1222 ESFIALYTDKLEYVHTFMNCFPRLEKFIINQDNDPKSQLOQCCLTLRT--EGKEPD 1280
831 EAFGLGALYVDKGLLYCEQCHVCLFRLQLFINQDNDPKSKLQCCCLTLRTMDGEPD 890
1281 IPLYKTLQTVGFSHARTYTVAVYFKGIGCGKGPSIQQAEMGAAMDALKYN--FPQMA 1338
891 IPYKVVVEASGPTNTRVYKVVAVVFSKRLATSSGSSIQQAEMNAKQALENSRDLFPOLD 950
1339 HOKRFGIRKYRQEL---KEMRWEREHOREPDETEDIKK 1374
951 HOKRVIKSIKQTGNELDNDSRQHE-----EKIKR 983

RESULT 14
Q7FNE9 ANOGA PRELIMINARY; PRT; 836 AA.
AC Q7FNE9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000011622 (Fragment).
GN ORFNames=ENSANG00000009133;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: the sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008964; EAA12298.3; -, Genomic_DNA.
```

```
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003725; F:double-stranded RNA binding; IEA.
DR GO: GO:0004519; F:endonuclease activity; IEA.
DR GO: GO:0004525; F:ribonuclease III activity; IEA.
DR GO: GO:0006396; P:RNA processing; IEA.
DR InterPro: IPR001159; DS RBD.
DR InterPro: IPR000999; RNase_III.
DR Pfam: PF00035; dcmr_1.
DR Pfam: PF00636; Ribonuclease_3; 2.
DR SMART: SM00358; DSRM; 1.
DR SMART: SM00535; RIBOC; 2.
DR PROSITE: PS00137; DS RBD; 1.
DR PROSITE: PS00517; RNase_3_1; 1.
DR PROSITE: PS00142; RNase_3_2; 2.
FT NON_TER 836 836
SQ SEQUENCE 836 AA; 96229 MW; 8D3E0BD04953877D CRC64;

Query Match 35.8%; Score 2683.5; DB 2; Length 836;
Best Local Similarity 59.2%; Pred. No. 7e-123;
Matches 499; Conservative 143; Mismatches 182; Indels 19; Gaps 8;

QY 514 HPDLHDELWYNDPGQNDGPKCKSAKARTGIRHSIYPGEBAIKPCRPMTNNAKRLPH 573
DB 1 HPLRLHVLWYNDGEMNDGPKCRCSARSRTGIRHGKYPGEGFPKCPINSNNADKLH 60

QY 574 YRITVSPPTNPLDRPTVIBYDHEIYIFEGFSMPAHAPLNIPLCKVIRFNIDYTHFIE 633
DB 61 YRITISPTTNFTKTPTIKHDKHDFEFLFEGLFSLAHEPIGELPTCKVIRFNIEVTL 120

QY 634 EMWPEFCVKGLFSLFLFRDILELDYDNLKGLPDESPCCPRFHFMPFRVFLPDGG 693
DB 121 EQPENFTIRELNFORYLFRELELVDTVQ-PSGSGEDSSCPYHFLFRFRVRLPDNG 179

QY 694 KEVLSHQIILLYLRCSKALVPBEEIANLQWELEWQKVAECCKGMVITNPGTKPSSVR 753
DB 180 KEVLAMSEVLRYLLDNSGVLPPDMLKEMDMQSEWQDVYDVYKGVNSVPGMKCSVR 239

QY 754 IDQLREQFNPVITPPIIVHFGIRPAQLSAGDPOYQKLWKSIVKLRHLANSKPKVQT 813
DB 240 VQQLDRNV--GDVPEANAIDENG-----YAGNPEYQKAWREYIKFRHLIANNKSPSE 290

QY 814 DKQKLAQREALKQIRQKNTMRREVTVELSSQCFWKTGIRSDVCOHAMMLPVLTHIRYH 873
DB 291 DRKLEAKENRLEDMRQGRMKNITAVSAKAFHRTGIMCDVMQVAMLPVLTGHLRFH 350

QY 874 QCLMHLDKLIGYFQDRLCLLQAMTHPSHHLNFGMNPDHARNSLSCNGIRQPKYGRKVH 933
DB 351 RSLNVLEIYGYFTYRYTLQALTHPSYKENFGTHPDHARNSLTNCGIRQPEYGRKH 410

QY 934 HMMRKKGINTLINMSRLGQDDPTPSRINHNRLRFLGDAVVEPLTSHVLYLPPSLRE 993
DB 411 YMNTRKGINTLISIMSRFKEHETDSNITHNRLRFLGDAVVEFITSIHLFHMFPDLSE 470

QY 994 GGLATYRTAIVQNHMLAKKLELDPFMYLHAGPDLCHESDLRHAMANCFEALIGAVYL 1053
DB 471 GGLATYRTAIVQNHMLAVLAKKLELDFEFLMYLHAGSDLCHELELRHALANCFEALMGALL 530

QY 1054 EGSLEBAKOLFGRLFNDDPD-LREVLNYPVLPQLQEPNTDRLQLTETSPLVQLTEREE 1112
DB 531 DGGIEVADRVAFVALFQEDDTLGIWNTYPSHPLEQEPGLGRKHIDSEMLKTLTRFED 590

QY 1113 AIGVIFTHVRLARAFLTRVGFNHLTGLHGNQMBFLGDSIMQVATEYLFIFHPDHEG 1172
DB 591 SIGVQFNHILLARAFTRDSIGFTNLTLGNSQRLFLGDTVLQLICSEVLYRHPPEHGG 650

QY 1173 HLTLLRSSLVNNTQAKVAELGMQEIYAI-TNDKTRPVGLRTKTLADLLESFIAALYTD 1231
DB 651 HLSLLRSSLVNNTQAVVCDLGMTOYAVYSNPK----ADLTKDRADLLEAFGLALYVD 706

QY 1232 KDLVYHTFMVCFPRLEKEFILLQNDNDPKSLQOCCLTLRT-EGKEPDIPLYKTLQTV 1290
DB 707 KGLEICEMFCHVCLFRLQDFINQDNDPKSLQOCCLTLRTMDGGBDIPVYKVIECT 766

QY 1291 GPSHARTYTVAVYFKGERTGCGKPSIQQAEMGAAMDALE--KYNFPQMAHQRFIGRKY 1348
DB 767 GPTNTVYVAVYFRGKRLACADGHSIQQAEMNAKQALENSKDLFPQLDQHRVIAQSL 826

QY 1349 ROE 1351
DB 827 KRQ 829

RESULT 15
Q59FF9 HUMAN
ID Q59FF9_HUMAN PRELIMINARY; PRT; 432 AA.
AC Q59FF9;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Nuclear RNase III Drosha variant (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209501; BAD92738.1; -; mRNA.
FT NON_TER 1
SQ SEQUENCE 432 AA; 50659 MW; E5D811B6DD61251A CRC64;

Query Match 30.6%; Score 2298; DB 2; Length 432;
Best Local Similarity 99.8%; Pred. No. 2.1e-104;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 556 EAIKPCRPMTNNAKRLFHYRITVSPPTNFLTDRPTVIEYDDHEIYIFEGFSMPAHAPLNI 615
DB 7 QAIKPCRPMTNNAKRLFHYRITVSPPTNFLTDRPTVIEYDDHEIYIFEGFSMPAHAPLNI 66

QY 616 PLCKVIRFNIDYTHIPIEMMPENFCVKGLELFLFRDILELDYDNLKGLPDESPPC 675
DB 67 PLCKVIRFNIDYTHIPIEMMPENFCVKGLELFLFRDILELDYDNLKGLPDESPPC 126

QY 676 CRRFHPMPFRVFLPDGGKEVLSMHQIILLYLRCSKALVPBEEIANLQWELEWQKVAE 735
DB 127 CRRFHPMPFRVFLPDGGKEVLSMHQIILLYLRCSKALVPBEEIANLQWELEWQKVAE 186

QY 736 ECKGMIVTNPGTKPSSVRIDQLDRSQFNPDVITFPIIVHFGIRPAQLSYAGDPQYQKLMK 795
DB 187 ECKGMIVTNPGTKPSSVRIDQLDRSQFNPDVITFPIIVHFGIRPAQLSYAGDPQYQKLMK 246

QY 796 SYVKLRHLANSKPKVQTKDKLAOREEALQKIRQKNTMRREVTVELSSQGFWKTGIRSD 855
DB 247 SYVKLRHLANSKPKVQTKDKLAOREEALQKIRQKNTMRREVTVELSSQGFWKTGIRSD 306

QY 856 VCOHAMMLPVLTHIRYHOCMLHLDKLGITYTFQDRLCLLQAMTHPSHHLNFGMNPDHARN 915
DB 307 VCOHAMMLPVLTHIRYHOCMLHLDKLGITYTFQDRLCLLQAMTHPSHHLNFGMNPDHARN 366

QY 916 SILSNCGIRQPKYGRKRVHMMRKKGINTLINMSRLGQDDPTPSRINHNRLRFLGDAV 975
DB 367 SILSNCGIRQPKYGRKRVHMMRKKGINTLINMSRLGQDDPTPSRINHNRLRFLGDAV 426

QY 976 VSEFLT 980
DB 427 VSEFLT 431

Search completed: December 24, 2005, 00:58:47
Job time : 277 secs
```
